universität wien

DIPLOMARBEIT

Titel der Diplomarbeit

Mimicking Stochastic Processes
Martingales & Marginals

Verfasser

Mag.phil. Dr.phil. Friedrich Penkner

angestrebter akademischer Grad

Magister der Naturwissenschaften (Mag.rer.nat)

Wien, im Dezember 2010

Studienkennzahl lt. Studienblatt:	A 405
Studienrichtung lt. Studienblatt:	Mathematik
Betreuer:	Dipl.Ing. Dr. Mathias Beiglböck

Der allerliebsten Zwetschke,
Helena-Lotta

Abstract

The problem of mimicking stochastic processes has become popular over the last 15 years, pre-eminently in the context of mathematical finance. *Mimicking* here roughly means: given a stochastic process X, find another process $\widetilde{X}$ which is equal to the original process in all one dimensional marginals. In math-finance this is of particular interest since there is a one-to-one correspondence between the 1-d marginals of an asset price process and the prices of European call options on this asset. If the mimicking process is constructed as a Markov process, the procedure is called *Markovian projection*.

In Chapter 1 we give a brief summary of stochastic calculus, including martingale and Markov theory and some theory of stochastic differential equations.

Chapter 2 is concerned with mimicking the most important continuous time process with continuous state space: Brownian motion. We discuss several constructions of fake Brownian motions, continuous and discontinuous ones. Furthermore, we characterize (linear) Brownian motion as the unique continuous strong Markov martingale having Brownian marginals.

Chapter 3 enlarges the scope of processes to be mimicked to real valued martingales and k-dimensional Itō-processes. Both in the case of real valued martingales and in the case of k-dimensional Itō-processes, we identify the classes of processes for which the Markovian projection is well defined.

Zusammenfassung

Besonders in den letzten 15 Jahren hat das Interesse an *Mimicking* von stochastischen Prozessen immens zugenommen, vor allem im Bereich Finanzmathematik. *Mimicking* bedeutet im Wesentlichen, für einen gegebenen stochastischen Prozess X einen anderen Prozess $\widetilde{X}$ zu finden, der in allen eindimensionalen Randverteilungen gleich dem ursprünglichen Prozess ist. In der Finanzmathematik ist das nicht zuletzt deswegen von Interesse, da die Familie der eindimensionalen Randverteilungen des Preis-Prozesses einer Aktie eineindeutig der Menge aller Preise einer European call option auf diese Aktie entspricht. Ist der Mimicking-Prozess Markov, bezeichnet man ihn als Markov-Projektion des ursprünglichen Prozesses.

In Kapitel 1 wird ein kurzer Abriss der stochastischen Analysis gegeben, vor allem der Theorie der Martingale, Markovprozesse und stochastischen Differentialgleichungen.

Kapitel 2 beschäftigt sich mit Mimicking-Resultaten zur Brownschen Bewegung, dem wichtigsten stetigen stochastischen Prozess mit kontinuierlichem Zustandsraum. Es werden verschiedene Konstruktionen von *fake Brownian motions* vorgestellt, die stetige wie auch nichtstetige Prozesse liefern. Schließlich wird die Brownsche Bewegung als eindeutiges, stetiges, starkes Markov-Martingal charakterisiert, das die entsprechenden Randverteilungen aufweist.

In Kapitel 3 werden Mimicking Resultate zu Martingalen und Itō-Prozessen im Allgemeinen diskutiert. Sowohl für reellwertige Martingale als auch für k-dimensionale Itō-Prozesse beschreiben wir schließlich jeweils eine Klasse von Prozessen für die die Markov-Projektion wohldefiniert ist.

Acknowledgements

There are many people whom I owe a lot. In the field of mathematics and in all other fields. Mathematics is easier than life, so I will start with the mathematical acknowledgements.

First of all, I would like to thank my advisor, Mathias Beiglböck, who suggested the topic of mimicking to me. In particular, I would like to thank him for the permanent and enduring support. During the last 11 months he always had a sympathetic ear for problems and suggestions, but rarely the time to discuss them. At the end of the day, the work turned out to be really interesting and challenging. Many thanks to you. Stay the way you are.

Secondly, I have to express my deep gratitude to Walter Schachermayer, who was not only the *éminence grise* behind the scenes of this diploma thesis but also encouraged me to dive a little into recent research. In particular, the open-minded and collegial athmosphere at the conference in honor of his 60th birthday was deeply impressive. (Even for me, participating just as a student.) Above all, it is due to him and Mathias that I try to go for more.

Mathematically, I also have to thank my little daughter Helena-Lotta for the really deep impact her upcoming birth had on my efforts to write a thesis.

But there also exists a non-mathematical side of life, which is why I would like to thank my families: My parents in Steinbach, who always respected what I was doing, even if they didn't like or understand it, and who supported me nevertheless; my beloved brothers Stefan and Bernhard and my dearest sister Angela, who is one of my supporting pillars.

And, last but not least, my own, even more beloved, little family: Sylvia and *Zwetschke* Helena-Lotta. It is them who constitute almost all my life and I am deeply grateful for the experience to share everyday life with two such wonderful girls.

Danksagung

Danke.

P.S.: Das richtet sich vor allem an jene, die im englischsprachigen Teil nicht erwähnt wurden. Also, an alle Freunde und Freundinnen, ohne die das Leben nur halb so lustig wäre, z.B. Pfaff, Luki, meine Adoptivschwestern Chri und Edith, Manuel, Gotti, Franzi, Kathi, Puschel, Max, Simon, Sandra, Hannes, Ana, Robert, Arno, Lydia, Magi, Mario, Klausi, die Prostestanten, Aunty Gela, Aunty Joana, Romani, Doris, Esther, Michou, Lisa, Richard, Ralf, Rotsch, Sebastian, Sevi, Thomas, Bauli, Hubert, Ursula, Andrea, Antonio, Berni, Birgit, Antonia, Barbara, Vroni, Luise, David, Karo, Horst, Karin, Josh, Friederike, Geri, Clemens, Dani, Verena und alle, die ich aus Platzgründen schlicht und ergreifend nicht erwähnen kann.

P.P.S.: Meinen Eltern möchte ich auch auf Deutsch noch einmal danken, für alles und sowieso: Danke. Wirklich.

P.P.P.S.: Ah, ja. Mathias. Danke auch noch einmal für das Nichtverzweifeln an meinem zuweilen etwas eigenwilligen Englisch.

P.P.P.P.S.: Meine Mädels. Auch noch einmal auf Deutsch: Danke. I like you.

Contents

List of Notation

$\mathbb{N}$ .... the set of the natural numbers. $\mathbb{N} = \{1, 2, 3, \dots\}$.

$\mathbb{R}$ .... the set of the real numbers.

$\mathbb{Q}$ .... the set of the rational numbers.

$\mathbb{R}^+$, $\mathbb{Q}^+$ .... the positive real/rational numbers: $\mathbb{R}^+ := [0, \infty)$, $\mathbb{Q}^+ := \mathbb{Q} \cap \mathbb{R}^+$.

$(\mathbb{R}^d)^{[0,\infty)}$ .... the set of functions from $[0, \infty)$ to $\mathbb{R}^d$.

$\mathcal{C}[0, \infty)^d$ .... the subspace of $(\mathbb{R}^d)^{[0,\infty)}$ consisting of the continuous functions.

$\mathcal{D}[0, \infty)$ .... the subspace of $\mathbb{R}^{[0,\infty)}$ consisting of the *cádlág* functions.

$u_t, u_{x_i}, u_{x_i x_j}$ ...... the partial derivatives $\frac{\partial u}{\partial t}$, $\frac{\partial u}{\partial x_i}$, $\frac{\partial^2 u}{\partial x_i \partial x_j}$.

$\mathcal{B}(U)$ .... the Borel-σ-algebra on the topological space U.

$\sigma(\mathcal{G})$ .... the smallest σ-algebra containing the collection of sets $\mathcal{G}$.

$\sigma(X_s)$ .... the smallest σ-algebra with respect to which the r.v. X_s is measurable.

$\sigma(X_s; 0 \leq s \leq t)$ .. the smallest σ-algebra with respect to which the r.v. X_s is measurable $\forall s \in [0, t]$.

$\mathcal{F}_t^X$ .... $\mathcal{F}_t^X := \sigma(X_s; 0 \leq s \leq t)$, $\mathcal{F}_\infty := \sigma(\bigcup_{t \geq 0} \mathcal{F}_t)$

$\mathcal{F}_{t+}$ .... $\mathcal{F}_{t+} := \bigcap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon}$

$\mathbb{P}^x$ .... the probability measure corresponding to Brownian motion or a Markov process with initial position $x \in \mathbb{R}^d$.

$\mathbb{E}^x$ .... denotes the expectation with respect to the probability measure $\mathbb{P}^x$.

$\mathbb{P}^\mu$ .... the probability measure corresponding to Brownian motion or a Markov process with initial distribution μ.

$\mathbb{P}_X$ .... the push forward measure unter a measurable function X. (E.g. the distribution of a random variable.)

$\mathcal{P}_X^f$ .... the family of all finite dimensional distributions of a process X.

$\mathcal{N}(\mu, \sigma^2)$ .... the Gaussian distribution with mean μ and variance σ^2.

0 Introduction – Preliminaries

0.1 Motivation. Why 1-d marginals?

One of the key concepts in identifying stochastic processes is the notion of identity in the *finite dimensional distributions*. On the one hand, most of the properties of a stochastic process $(X_t)_{t\geq 0}$ are determined via its family of finite dimensional distributions, on the other hand, using Kolmogorov's extension criterion, a stochastic process can be defined via a family of finite dimensional distributions satisfying two mild consistency conditions. The uniqueness of the so defined measure on the path space gives a very powerful concept of uniqueness, since for two processes' being equal in the f.d.d.s, it's not even necessary to be defined on the same probability space. These results have been well known for decades. The aim of this thesis is to consider processes that are equal just in all one dimensional marginal distributions and to study the interplay of the equality in these marginals and other properties of the processes, like being continuous, a martingale or Markov.

But what is the use of weakening the concept of identity in the f.d.d.s and considering just identity in the one dimensional marginals (dropping all joint distributions)? Essentially, there are two and a half reasons, why this could be of interest.

Reason (1): *Construction – Fitting.* Given a sequence $(\mu_n)_{n\geq 0}$ (or a continuum $(\mu_t)_{t\geq 0}$) of probability measures, say, on $\mathbb{R}$, one could be interested in finding reasonable stochastic processes that match these marginals. In mathematical finance, for example, the one dimensional marginals of a spot price process are given through the prices of European call options for all strikes and maturities. Using a local volatility model it is possible to fit a unique process to these marginals.

Reason (2): *Mimicking*: Given a badly behaved stochastic process, one is often interested in finding a better behaved process like a Markov process or a martingale having the same one dimensional marginals. This is of particular interest in applications where only the one dimensional marginal distributions matter.

Reason ($\mathbf{2\frac{1}{2}}$): *Theoretical issues*: A certain theoretical – and practical – interest lies in the question, how *close* different processes with the same one dimensional marginals are, or, under which assumptions the one-dimensional marginals suffice to determine the whole process (i.e. in all f.d.d.s). Moreover, the question arises whether a process, considered as a measure on the path space, depends continuously on the set of marginals that have to be matched.

In particular, one dimensional marginals play a crucial role when *real world phenomena* are modeled via stochastic processes. Not least because empirical observations usually are made at certain times $t_1 < \cdots < t_n$ and not simultaneously at these n times. So joint distributions very often are not available and automatically the question arises whether the modeling process having these marginals is unique (in the sense of f.d.d.s).

0.2 A first example

Take, for instance, the most important continuous time process with continuous state space, i.e. Brownian motion. We want to find a process $(X_t)_{t\geq 0}$ having the same one dimensional marginals, i.e. $X_t \sim \mathcal{N}(0,t)$ for all $t > 0$. If no other properties are required the process below is a solution to this problem (cf. [FWY00], p.452).

$$X_t := \sqrt{t}\, N, \tag{0.2.1}$$

where N is a standard normal random variable on the real line, $N \sim \mathcal{N}(0,1)$.

Let us look at the properties of $(X_t)_{t\geq 0}$. By construction $X_t \overset{(law)}{=} B_t$ for all $t \geq 0$, moreover we see that all paths are continuous, like in the case of Brownian motion. Nevertheless the process (X_t) is highly different from Brownian Motion in various aspects.

First of all, the only randomness in (X_t) is the magnitude of the drift; all paths are scaled graphs of the square root, i.e. the law of the process (X_t) on the canonical path space $\Omega = \mathbb{R}^{[0,\infty)}$ is concentrated on the set $\widetilde{\Omega} = \{f : \mathbb{R}^+ \to \mathbb{R} \mid f(t) = a\sqrt{t}; a \in \mathbb{R}\} \subsetneq \mathcal{C}[0,\infty)$. Therefore the law of (X_t) cannot be the Wiener measure. Obviously the process is not a martingale since there is just a drift but no diffusion term, but $(X_t)_{t\geq 0}$ is Markov in a trivial sense. Finally, this process is of bounded variation, hence the quadratic variation is zero, indicating another difference to Brownian motion.

In order to get closer to Brownian motion we could look for a mimicking process having some more randomness. Hence, again we are looking for a continuous process (Y_t) with the same one dimensional marginals as Brownian motion but having nonzero quadratic variation. Let $(B_t)_{t\in[0,1]}$ be a Brownian motion on the unit interval and define (cf. [FWY00], p.452)

$$Y_t = \begin{cases} B_t & \text{for } t \leq \frac{1}{2} \\ B_{\frac{1}{2}} + (\sqrt{2}-1)B_{t-\frac{1}{2}} & \text{for } t > \frac{1}{2}. \end{cases} \tag{0.2.2}$$

Then $Y_t \overset{(law)}{=} B_t$ for all $t \in [0,1]$ and the quadratic variation is given by

$$d\langle Y\rangle_t = \begin{cases} dt & \text{for } t \leq \frac{1}{2} \\ (\sqrt{2}-1)^2 dt & \text{for } t > \frac{1}{2}. \end{cases} \tag{0.2.3}$$

However, Y is not a (continuous) semimartingale, since at $t = \frac{1}{2}$ there is a jump in the finite variation process, in particular Y cannot be a (local) martingale. But, it is a time inhomogeneous Markov process.

So, it should be kept in mind that there is a huge variety of processes having the same one dimensional marginals. But, depending on additionally required properties, we can make the mimicking process look more and more like the original process. For instance, a sample path of $(Y_t)_{t\in[0,1]}$ cannot be distinguished from a Brownian sample path as easily as a path of $(X_t)_{t\geq 0}$.

0.3 A second example

As indicated in *reason* (1), often we are given a (more or less) arbitrary family of probability measures $(\mu_t)_{t\geq 0}$ and we try to find a reasonable process having these marginals. For instance, we could look for a *diffusion*, i.e. a solution of a stochastic differential equation, having the given marginals. In the case of local volatility models, this was done by Bruno Dupire in 1994. Dupire reconstructed the spot price process $(S_t)_{t\geq 0}$ of an asset with the help of the family of one dimensional marginals (see [Dup97]). The key observation was that, if the (risk neutral) spot price follows a one dimensional diffusion, then knowing the prices of a European call option on the asset for all strikes and maturities is equivalent to the knowledge of the one dimensional distributions $\mu_t(dx) = \mathbb{P}(S_t \in dx)$ of the spot price-process.

Proposition 0.3.1. *Let $(X_t)_{t\geq 0}$ be real valued and let $(\mu_t)_{t\geq 0}$ be the family of its one dimensional marginal distributions. Then the function $C\colon \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$; $C(t,k) := \mathbb{E}[(X_t - k)^+]$ uniquely determines the family $(\mu_t)_{t\geq 0}$ and vice versa.*

Proof. First, assume that we are given the family (μ_t). Then clearly for all $t \geq 0$ and all $k \in \mathbb{R}$ the values $C(t,k)$ are uniquely determined.

The other direction we just show in the special case where the $\mu_t(dy)$ admit densities $p(t,y)dy$, but it also holds in the general case. In the special case

$$\begin{aligned} C(t,k) = \mathbb{E}[(X_t - k)^+] &= \int_{\mathbb{R}} (y-k)^+ \, d\mu_t(dy) = \int_{\mathbb{R}} (y-k)^+ \, p(t,y) \, dy = \\ &= \int_k^\infty (y-k) p(t,y) \, dy. \end{aligned}$$

We obtain, differentiating twice w.r.t. k,

$$\begin{aligned} C_k(t,k) &= -\int_k^\infty p(t,y) \, dy \\ C_{kk}(t,k) &= p(t,k). \end{aligned}$$

So, for each $t \geq 0$ we get a density p in k for the value of X_t. $\square$

Dupire now considered the Black-Scholes model, assuming $r = 0$ for the interest rate, but replaced the constant volatility σ by a local volatility process $\sigma(t, S_t)$. Under the risk neutral (martingale-)measure $\widetilde{\mathbb{P}}$ the spot price process then follows the equation

$$\frac{dS_t}{S_t} = \sigma(t, S_t) \, d\widetilde{B}_t \tag{0.3.1}$$

and the prices of the European calls with strikes K and maturities T are given via

$$C(T,K) = \widetilde{\mathbb{E}}[(S_T - K)^+]. \tag{0.3.2}$$

As we saw in Proposition 0.3.1, the knowledge of all these prices is equivalent to the knowledge of the one-dimensional marginals. With the help of the Fokker-Planck equation Dupire recovered the instantaneous volatility

$$\sigma(T,K)^2 = \frac{C_T(T,K)}{\frac{1}{2}K^2 C_{KK}(T,K)} \tag{0.3.3}$$

and hence the whole (unique) spot price process. (The factor K^2 in the denominator is due to the use of geometric Brownian motion.)

The so constructed diffusion is unique, as Dupire states in [Dup97], because the spot price process is supposed to be a martingale. However, what Dupire does not tell, is *how* the martingale property ensures uniqueness. One could argue that, since the one-dimensional market model used by Dupire is complete, by the (second) Fundamental Theorem of Asset Pricing (cf. [Shr04], p.232), there exists a unique risk free (martingale) measure $\widetilde{\mathbb{P}}$, such that under $\widetilde{\mathbb{P}}$ the (discounted) spot price process is a martingale and equals (0.3.1). So there is some indication of uniqueness. Although there still could be a martingale, having these marginals but different joint distributions. Why in this special setting the one-dimensional marginals suffice to determine the whole process will be answered in Chapter 3.

0.4 Outline and Summary

In Chapter 1 we give a brief introduction to stochastic calculus, mainly without proofs, since the used machinery is standard nowadays.

In Chapter 2 we discuss the existing mimicking results for Brownian motion in some detail and approximate, in a certain sense, the standard Brownian motion by processes which are increasingly better behaved. In requesting more regularity, we finally get close to Brownian motion itself.

In Chapter 3 we consider martingales and marginals. We present three rather theoretical results of Lowther concerning the unique fitting of marginals to martingales within a certain class of strong Markov martingales. Finally we discuss the milestone-result of Gyöngy which allows to mimick continuous Itō-processes via SDEs.

To put it in a nutshell, the main aim of the following pages is to provide a path through the jungle of existing mimicking and fitting results and to give an overview of what is currently at stake with respect to the question of finding processes which have given one-dimensional marginals.

1 Stochastic Calculus

First we recall some basic notions and mention some of the important existence and uniqueness theorems. On the one hand to clarify the setting in which we are working, on the other to fix notation. For a complete and systematic description of stochastic calculus and all the subtleties see [RY99], [KS88] or [Pro04].

1.1 Stochastic processes – Canonical versions

We always assume as given a complete, filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t\geq 0})$ satisfying the *usual hypotheses*. A filtration is an increasing family of σ-algebras $(\mathcal{F}_t)_{t\in T}$ s.t. $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $s \leq t$. T denotes the parameter set. As usual T will be regarded as time and we will almost all the time deal with $T = [a, b]$; $a, b \in \mathbb{R}$ or $T = \mathbb{R}^+$.

Definition 1.1.1. (cf. [Pro04], p.3) A complete filtered probability space is said to satisfy the *usual hypotheses* if:

(i) $\mathcal{F}_0$ contains all the $\mathbb{P}$-null sets.

(ii) $\mathcal{F}_t = \mathcal{F}_{t+} := \bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon}$ for all $t \in [0, \infty)$; i.e. $(\mathcal{F}_t)_{t\geq 0}$ is right continuous.

Although there are several ways to view stochastic processes, we define a stochastic process as a parametrized collection of random variables.

Definition 1.1.2. Let T be a parameter set and $(E, \mathcal{E})$ a measurable space. A stochastic process indexed by T is a family of measurable mappings X_t, $t \in T$ from a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to the state space $(E, \mathcal{E})$.

As we consider only real- or $\mathbb{R}^d$-valued processes, from now on the state space always will be $E = \mathbb{R}$, $E = \mathbb{R}^d$ (or a Borel subset) equipped with the Borel σ-algebra: $\mathcal{E} = \mathcal{B}(\mathbb{R})$, $\mathcal{E} = \mathcal{B}(\mathbb{R}^d)$ respectively.

For fixed $\omega \in \Omega$ the mappings $t \mapsto X_t(\omega)$, $t \in T$, are called *sample paths*, which we usually will require to be continuous. A *process* is said to be continuous if almost all paths are continuous.

Definition 1.1.3. Let $(E, \mathcal{E})$ be a topological space, endowed with its Borel σ-algebra. A process with state space $(E, \mathcal{E})$ is said to be a.s. continuous if for almost all ω the function $t \to X_t(\omega)$ is continuous.

We will need notions to compare resp. identify stochastic processes in certain ways.

Definition 1.1.4. Let $(X_t)_{t\in T}$ and $(X_t)'_{t\in T}$ be stochastic processes, defined on probability spaces $(\Omega, \mathcal{F}, \mathbb{P})$ and $(\Omega', \mathcal{F}', \mathbb{P}')$ respectively.

(i) They are said to be *versions* of each other :⇔

$$\begin{aligned}&\forall A_1, \dots, A_n \subset E, \quad A_i \in \mathcal{B}(E), \quad t_1, \dots, t_n \in T: \\ &\mathbb{P}(X_{t_1} \in A_1, \dots, X_{t_n} \in A_n) = \mathbb{P}(X'_{t_1} \in A_1, \dots, X'_{t_n} \in A_n).\end{aligned}$$

In this case we also say that the *finite dimensional distributions* of $(X_t)_{t\in T}$ and $(X_t)'_{t\in T}$ coincide. We will denote this by $(X_t)_{t\in T} \overset{(fdd)}{=} (X_t)'_{t\in T}$.

(ii) $(X_t)_{t\in T}$ and $(X_t)'_{t\in T}$ are said to be *modifications* of each other :⇔

$$(\Omega, \mathcal{F}, \mathbb{P}) = (\Omega', \mathcal{F}', \mathbb{P}'), \text{ and } \forall t \in T\colon \mathbb{P}(X_t = X'_t) = 1.$$

(iii) $(X_t)_{t\in T}$ and $(X_t)'_{t\in T}$ are said to be *indistinguishable* :⇔

$$(\Omega, \mathcal{F}, \mathbb{P}) = (\Omega', \mathcal{F}', \mathbb{P}'), \text{ and } \mathbb{P}(X_t = X'_t;\, \forall t \in T) = 1.$$

Remark 1.1.5. In general: (iii) $\implies$ (ii) $\implies$ (i). The notion *modification* is weaker than indistinguishability in the sense that for each t a (probably different) null set $N_t := \{\omega | X_t(\omega) \neq X'_t(\omega)\}$ exists, where the processes differ. Since any real interval $T = [0, s]$, $s > 0$ is uncountable, the set $N := \bigcup_{t\in T} N_t$ could have any probability between 0 and 1, or be even non-measurable. If, however, $(X_t)_{t\in T}$ and $(X_t)'_{t\in T}$ are *indistinguishable*, then there is only one null set N, and for $\omega \notin N$: $X_t(\omega) = X'_t(\omega)$ for all $t \in T$. This means precisely that the sample paths really are the same outside a null set. (cf. [Pro04], p.4)

Lemma 1.1.6. *If $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ have continuous paths and are modifications of each other, then they are indistinguishable.*

Proof. First, note that $\mathbb{R}^d$-valued continuous functions are uniquely determined by their values on a dense subset. We set $D = [0, \infty) \cap \mathbb{Q}$. Clearly D is dense in $\mathbb{R}^+$ and countable. We denote by N_t the null set where $X_t \neq Y_t$. Then

$$0 = \mathbb{P}\left(\bigcup_{t \in D} N_t\right) = \mathbb{P}\left(\exists t \in D \colon X_t \neq Y_t\right) = 1 - \mathbb{P}(\forall t \in D \colon X_t = Y_t).$$

□

Another way to denote the identity in (i) is to define the family of the f.d.d.s for a given process X, i.e. the set of the push forward measures on all k-dimensional products of the state space.

Definition 1.1.7. Let $X = (X_t)_{t \in T}$ be an $\mathbb{R}^n$ valued stochastic process, $t_i \in T \subseteq \mathbb{R}^+$. The family of *finite dimensional distributions* of X is the set

$$\mathcal{P}_X^f := \{\mathbb{P}_{X;t_1,\dots,t_k} \,|\, k \in \mathbb{N},\, t_i \in T\},$$

where the probability measures $\mathbb{P}_{X;t_1,\dots,t_k}$ on $\mathbb{R}^{nk}$ are defined via

$$\mathbb{P}_{X;t_1,\dots,t_k}(A_1 \times \cdots \times A_k) := \mathbb{P}(X_{t_1} \in A_1, \dots, X_{t_k} \in A_k), \quad k \in \mathbb{N}, A_i \in \mathcal{B}(\mathbb{R}^n).$$

Hence $(X_t)_{t \in T}$ and $(Y_t)_{t \in T}$ are versions of each other if and only if $\mathcal{P}_X^f = \mathcal{P}_Y^f$.

Most of the time we are precisely interested in the concept of identity given through *versions*, since the Kolmogorov extension theorem allows to define a stochastic process in terms of the finite dimensional distributions, which means that, given a family of finite dimensional distributions satisfying a *permutation property* and a *restriction property*, there exists a process having these finite dimensional distributions.

Theorem 1. *(Kolmogorov's Extension criterion, cf. [Øks98], p.11)*
Let $T \subseteq \mathbb{R}$. Let, for all $t_1, \dots, t_k \in T$ and $k \in \mathbb{N}$, $\nu_{t_1,\dots,t_k}$ be probability measures on $\mathbb{R}^{nk}$ s.t.

$$\nu_{t_{\pi(1)},\dots,t_{\pi(k)}}(A_1 \times \cdots \times A_k) = \nu_{t_1,\dots,t_k}(A_{\pi^{-1}(1)} \times \cdots \times A_{\pi^{-1}(k)}) \tag{1.1.1}$$

for all permutations π on $\{1, \dots, k\}$ and

$$\nu_{t_1,\dots,t_k}(A_1 \times \cdots \times A_k) = \nu_{t_1,\dots,t_k,t_{k+1},\dots,t_{k+m}}(A_1 \times \cdots \times A_k \times \mathbb{R}^n \times \cdots \mathbb{R}^n) \tag{1.1.2}$$

for all $m \in \mathbb{N}$, where the set on the r.h.s. has a total of $k+m$ factors.

Then there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a stochastic process $(X_t)_{t\in T}$, $X_t\colon \Omega \to \mathbb{R}^n$, s.t.

$$\nu_{t_1,\dots,t_k}(A_1 \times \cdots \times A_k) = \mathbb{P}(X_{t_1} \in A_1, \dots, X_{t_k} \in A_k)$$

for all $t_i \in T$, $k \in \mathbb{N}$ and $A_i \in \mathcal{B}(\mathbb{R}^n)$.

Now, given the existence of an (abstract) probability space Ω and a process $(X_t)_{t\geq 0}$ we are looking for some concrete probability space and a canonical version on it. Having in mind that, for each fixed $\omega \in \Omega$, we get a sample path $(t \mapsto X_t(\omega))$ we define

$$\Phi\colon \Omega \to (\mathbb{R}^n)^T, \quad \Phi(\omega) := (t \mapsto X_t(\omega)) \tag{1.1.3}$$
$$\Phi(\omega)(t) := X_t(\omega). \tag{1.1.4}$$

For $f \in (\mathbb{R}^n)^T$ we define the evaluation-, or coordinate functions

$$Y_{t_0}(f) := f(t_0), \quad Y_{t_0}(\Phi(\omega)) = \Phi(\omega)(t_0) = X_{t_0}(\omega) \tag{1.1.5}$$

which, for fixed t, are continuous w.r.t. the topological spaces $((\mathbb{R}^n)^T, \|\cdot\|_\infty)$ and $\mathbb{R}^n$, hence Borel-measurable. The image of $\mathbb{P}$ under Φ,

$$\mathbb{P}_X := \Phi(\mathbb{P}), \quad \mathbb{P}_X(B) := \mathbb{P}(\Phi^{-1}(B)), \quad B \in \mathcal{B}(\mathbb{R}^n)^T,$$

together with the path space defines a probability space. The process $(Y_t)_{t\in T}$ on this space is called the *coordinate process*. By construction we get the following

Proposition 1.1.8. *The coordinate process $(Y_t)_{t\in T}$ on $((\mathbb{R}^n)^T, \mathcal{B}(\mathbb{R}^n)^T, \mathbb{P}_X)$ is a* version *of the process $(X_t)_{t\in T}$ on $(\Omega, \mathcal{F}, \mathbb{P})$.*

Definition 1.1.9. $(Y_t)_{t\in T}$ is called the *canonical version* of X. The probability measure $\mathbb{P}_X$ is called the law of X.

Remark 1.1.10. Identifying ω and $(t \mapsto X_t(\omega))$ and the respective probability spaces $(\Omega, \mathcal{F}, \mathbb{P})$ and $((\mathbb{R}^n)^T, \mathcal{B}(\mathbb{R}^n)^T, \mathbb{P}_X)$ we sometimes will denote the coordinate process itself by ω, i.e. $Y_t(\omega) = \omega(t)$.

Accordingly, defining a stochastic process amounts to defining a probability measure on the path space. Therefore the Kolmogorov extension criterion could equivalently have been stated as

Theorem 2. *In the setting of Theorem 1, i.e. given a family of probability measures* $\mathcal{P}$ *satisfying the* restriction property *and the* permutation property*, there exists a unique probability measure* $\mathbb{P}$ *on* $((\mathbb{R}^n)^T, \mathcal{B}(\mathbb{R}^n)^T)$ *s.t. for the coordinate process* Y*,* $\mathcal{P} = \mathcal{P}_Y^f$.

But, as indicated above, we are not interested in an arbitrary process given by the extension criterion; we look for a continuous version or modification, the existence of which is ensured by the following:

Theorem 3. *(Kolmogorov's continuity criterion, cf. [RY99], Theorem I.1.8)*
A real valued process X *for which there exist three constants* α*,* β*,* $C > 0$ *such that*

$$\mathbb{E}[|X_{t+h} - X_t|^\alpha] \leq Ch^{1+\beta} \tag{1.1.6}$$

for every t *and* h*, has a modification which is almost surely continuous.*

The reason, why we have spent a possibly undue amount of space, or time, to these rather basic considerations should be obvious from the topic of the exposition. In what follows we shall proceed a bit faster.

1.2 Brownian Motion

As an application of the above we construct the (linear or one dimensional) Brownian motion via its Gaussian marginal densities. I.e. we construct a probability measure on $(\Omega, \mathcal{F}) = (\mathbb{R}^{[0,\infty)}, \mathcal{B}(\mathbb{R})^{[0,\infty)})$ s.t. the process $(B_t)_{t\geq 0}$ defined via the coordinate process, $B_t(\omega) = \omega(t)$ satisfies properties (i) and (ii) below, hence is a *version* of the standard Brownian motion. Finally the continuity criterion shall provide us with a continuous modification of the canonical version.

Definition 1.2.1. Let $(B_t)_{t\geq 0}$ be a real valued stochastic process on $(\Omega, \mathcal{F}, \mathbb{P})$. $(B_t)_{t\geq 0}$ is called a *(standard) Brownian Motion* iff the following conditions hold:

(i) $B_t \sim \mathcal{N}(0, t)$.

(ii) $\forall t_0 \leq \ldots \leq t_k; t_i \in \mathbb{R}^+$ and for $0 \leq i \leq k-1$: $(B_{t_{i+1}} - B_{t_i})$ are independent r.v.s.

(iii) $(B_t)_{t\geq 0}$ is a.s. continuous.

Conditions (i) and (ii) mean that Brownian motion has stationary independent increments which are centered Gaussian. The cumulative distribution function of the random vector $(B_{t_1}, \ldots, B_{t_k})$ for $0 = t_0 \leq t_1 \leq \cdots \leq t_k$ therefore has to be

$$\begin{aligned} F_{t_1,\ldots,t_k}(x_1, \ldots, x_k) &= \\ &= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_k} p(t_1; 0, y_1) p(t_2 - t_1; y_2, y_1) \cdots \\ &\cdots p(t_k - t_{k-1}; y_{k-1}, y_k)\, dy_k \cdots dy_2\, dy_1, \end{aligned} \tag{1.2.1}$$

for $(x_1, \ldots, x_k) \in \mathbb{R}^k$, and p the Gaussian kernel

$$p(t; x, y) := \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-y)^2}{2t}}, \quad t > 0,\ x, y \in \mathbb{R}.$$

Since the above definition is equivalent to the statement that the increments $\{(B_{t_j} - B_{t_{j-1}}) \,|\, 1 \leq j \leq k\}$ are independent and $(B_{t_j} - B_{t_{j-1}}) \sim \mathcal{N}(0, t_j - t_{j-1})$, we set

$$\mathbb{P}_{B;\, t_1,\ldots,t_k}(A_1 \times \cdots \times A_k) := \int\limits_{A_1\times\cdots\times A_k} p(t_1; 0, y_1) \cdots p(t_k - t_{k-1}, y_k, y_{k-1})\, dy_k \cdots dy_1$$

for $A_i \in \mathcal{B}(\mathbb{R})$, $0 \leq t_1 \leq \cdots \leq t_k$, and extend this definition via permutation of the time indices to all finite sequences $(t_i)_{i\in\{1,\ldots,k\}}$, $t_i \in \mathbb{R}^+$. Since $\int_{\mathbb{R}} p(t; x, y)\, dy = 1$ for all $t \leq 0$ and $x \in \mathbb{R}$, we can add finitely many factors and the restriction property (1.1.2) in Theorem 1 is satisfied. (cf. [KS88], p.52-53; [Øks98], p.11-13). Hence we get

Corollary 1.2.2. *There exists a (standard) Brownian motion* $(B_t)_{t\geq 0}$.

Proof. Define $\mathcal{P}_B^f := \{\mathbb{P}_{B;\, t_1,\ldots,t_k} \,|\, k \in \mathbb{N},\ t_i \in \mathbb{R}^+\}$. Then, as shown above, $\mathcal{P}_B^f$ satisfies the Kolmogorov extension criterion, i.e. there exists a probability measure $\mathbb{P}$ on $(\mathbb{R}^{[0,\infty)}, \mathcal{B}(\mathbb{R}^{[0,\infty)}))$ under which the coordinate process $\omega(t) = B_t(\omega)$, $\omega \in \mathbb{R}^{[0,\infty)}$, has stationary, independent increments and $(B_t - B_s) \sim \mathcal{N}(0, t - s)$ for $0 \leq s \leq t$.

To get a continuous modification we choose $\alpha = 4$ in the Kolmogorov continuity criterion and obtain, using the formulas for the moments of normally distributed random variables,

$$\mathbb{E}[|B_{t+h} - B_t|^4] = 3h^2 = Ch^{1+1}.$$

I.e. the continuity criterion is satisfied with $\alpha = 4$, $\beta = 1$ and $C = 1$, hence there exists standard Brownian motion. □

Remark 1.2.3. Since we have by now shown the existence of a process that is a.s. continuous, or put otherwise: since we have a probability measure on $(\mathbb{R}^{[0,\infty)}, \mathcal{B}(\mathbb{R}^{[0,\infty)}))$ that gives measure 1 to the subspace of real valued continuous functions $\mathcal{C}[0,\infty) \subset \mathbb{R}^{[0,\infty)}$, we will deal mainly with $\Omega = \mathcal{C}[0,\infty)$ or $\mathcal{C}[0,\infty)^d$.

Remark 1.2.4. For any $x \in \mathbb{R}$, the process $B^x = x + (B_t)_{t\geq 0}$ denotes the Brownian motion started at x. For $A \in \mathcal{B}(\mathbb{R})$, $t \geq 0$: $\mathbb{P}(B_t^x \in A) = \int_A p(t;x,y)\,dy$, hence $\mathbb{E}[B_t^x] \sim \mathcal{N}(x,t)$. From the point of view of canonical versions, i.e., when considering a stochastic process as a probability measure on a pathspace, however, it's more natural not to define a new process for each x but a new measure $\mathbb{P}^x$ (cf. e.g. [Bas95], p.9). For the coordinate process

$$X_t(\omega) = \omega(t), \quad \omega \in \mathcal{C}[0,\infty)$$

and $\mathcal{F} = \sigma(X_s; s < \infty)$ define $\mathbb{P}^x$ on $(\Omega, \mathcal{F})$ via

$$\mathbb{P}^x(X \in A) := \mathbb{P}(x + B_t \in A), \quad x \in \mathbb{R},\ A \in \mathcal{F}. \tag{1.2.2}$$

and call the pair $(\mathbb{P}^x, (X_t)_{t\geq 0})$ the Brownian motion started at x.

Remark 1.2.5. All constructions above generalize easily to $\mathbb{R}^d$-valued Brownian motions. $X_t = (B_t^1, \dots, B_t^d)$ is a d-dimensional Brownian motion iff all $(B_t^i)_{t\geq 0}$, $i \in \{1, \dots, d\}$ are independent linear Brownian motions. Using the d-dimensional Gaussian densities it's also possible to construct directly a d-dimensional Brownian motion, starting at $x = (x_1, \dots, x_1)$. (cf. [Øks98], p.11-13). In particular Brownian motion is the canonical example of a Gaussian process.

Definition 1.2.6. An $\mathbb{R}^d$ valued stochastic process $(X_t)_{t\geq 0}$ is called a *Gaussian process* if for any $k \in \mathbb{N}$ and $0 \leq t_1 < t_2 < \cdots < t_k < \infty$ the random vector $(X_{t_1}, \dots, X_{t_k})$ has a joint normal distribution. If the distribution of $(X_{t+t_1}, \dots, X_{t+t_k})$ does not depend on t, the process is called *stationary*.

Remark 1.2.7. There are some other ways to construct Brownian motion:

(i) *Donsker's invariance principle*: This is the most intuitive way to construct BM although the analysis underlying this construction is by no means trivial. BM is defined as a limit of properly scaled and affinely interpolated random walks on the integers. First, one shows via the multivariate Central Limit Theorem that the f.d.d.s of the random walk converge to the f.d.d.s of Brownian motion if the step size converges to zero. Secondly, one proves that the sequence of these laws is uniformly tight on $\mathcal{C}[0,\infty)$ (endowed with a somewhat weaker metric than $\|\cdot\|_\infty$). This shows that the random walk, considered as a path-valued random variable, converges weakly to Brownian motion. (Cf. [Shr04], p.83ff for a rather informational account, or [KS88], Section 2.4 for a full description.)

(ii) *Haar functions*: It is also possible, to construct Brownian motion on the compact unit interval via a Fourier series with random coefficients. Take $L^2([0,1],\lambda)$, where λ denotes the Lebesgue measure. Using the usual inner product $\langle f,g\rangle = \int_0^1 f(x)\overline{g}(x)\,d\,\lambda(x)$, the Haar functions φ_{ij} (cf. [KS88], p.58) form an orthonormal basis of $L^2([0,1],\lambda)$. Let U_{ij} be an iid sequence of standard normal r.v.s, define $\psi_{ij}(t) = \int_0^t \varphi_{i,j}(s)\,ds$ and $V_i(t) = \sum_{j=1}^{2^{i-1}} U_{ij}\psi_{ij}(t)$, then

$$B(t) := \sum_{i=0}^{\infty} V_i(t). \tag{1.2.3}$$

The sum converges uniformly in t and is a standard Brownian motion. (cf. [Bas95], p.11) The advantage of this construction is that by construction the paths are continuous, since the $\psi_{ij}(t)$ are continuous and the series converges uniformly on a compact set.

The reason why we adopted the construction via the Kolmogorov criterion lies in the fact that this construction gives the clearest view of the role of the distribution(s) and how certain processes are related via its f.d.d.s.

1.3 Martingales

In this section we give a brief summary of martingale theory, mainly taken from [RY99], Sections I.4 and II.1.

Definition 1.3.1. A process $(X_t)_{t\geq 0}$ on $(\Omega,\mathcal{F},\mathbb{P};(\mathcal{F}_t)_{t\geq 0})$ is *adapted* to the filtration $(\mathcal{F}_t)_{t\geq 0}$ if X_t is $\mathcal{F}_t$-measurable for each $t\geq 0$.

Definition 1.3.2. Given a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t\geq 0})$ a real valued process $(X_t)_{t\geq 0}$ adapted to $(\mathcal{F}_t)_{t\geq 0}$ is called a *submartingale* (w.r.t. $(\mathcal{F}_t)_{t\geq 0}$) if:

(i) $\mathbb{E}[|X_t|] < \infty$ for every $t \geq 0$
(ii) $\mathbb{E}[X_t \,|\, \mathcal{F}_s] \geq X_s$ a.s. for every pair s, t s.t. $s \leq t$.

A process X s.t. $-X$ is a submartingale is called a *supermartingale* and if X is both a sub- and a supermartingale, it's called a *martingale*.

Example 1.3.3. As is commonly known $(B_t)_{t\geq 0}$ is a martingale (w.r.t. $\mathcal{F}_t^B = \sigma(B_s; s \leq t)$), and so are $(B_t)^2_{t\geq 0} - t$ and $M_t^\alpha := \exp(\alpha B_t - \frac{\alpha^2}{2}t)$, for $\alpha \in \mathbb{R}$. This is, in each case, an easy consequence of the independence of the increments, properties of conditional expectation and the centered Gaussian distribution. E.g. for $s \leq t$:

$$
\begin{aligned}
\mathbb{E}[B_t | \mathcal{F}_s] &= \mathbb{E}[(B_t - B_s) + B_s | \mathcal{F}_s] = \\
&= \mathbb{E}[B_t - B_s] + \mathbb{E}[B_s | \mathcal{F}_s] = 0 + B_s = B_s \\
\mathbb{E}[B_t^2 - t | \mathcal{F}_s] &= \mathbb{E}[(B_t - B_s)^2 + 2B_t B_s - B_s^2 - t | \mathcal{F}_s] = \\
&= \mathbb{E}[(B_t - B_s)^2] + 2B_s \mathbb{E}[B_t | \mathcal{F}_s] + B_s^2 - t = \\
&= t - s + 2B_s^2 - B_s^2 - t = B_s^2 - s
\end{aligned}
$$

Remark 1.3.4. $\mathbb{E}[|B_t|^k] < \infty$ for all $t \geq 0$ and $k \in \mathbb{N}$ since all moments of normally distributed random variables exist. If not stated otherwise, we assume in the sequel of the exposition that all processes – especially (sub-, super-) martingales – as integrable, i.e. $X_t \in L^1(\Omega, \mathbb{P})$ for all t.

The conditional version of Jensen's inequality shows that submartingales arise naturally as functionals of martingales:

Proposition 1.3.5. *Let $(X_t)_{t\geq 0}$ be a real valued $(\mathcal{F}_t)$-martingale and $\psi\colon \mathbb{R} \to \mathbb{R}$ a convex function such that $\mathbb{E}[|\psi(X_t)|] < \infty$ for every t. Then $(\psi(X_t))_{t\geq 0}$ is an $(\mathcal{F}_t)$-submartingale.*

Proof. Let $s \leq t$, then:

$$
\mathbb{E}[\psi(X_t) | \mathcal{F}_s] \geq \psi(\mathbb{E}[X_t | \mathcal{F}_s]) = \psi(X_s). \tag{1.3.1}
$$

$\square$

So, given a martingale X, for instance $(|X_t|^p)_t$ for $p \geq 1$ is a submartingale (if $X_t \in L^p$), just to mention an important case (covering the modulus and the square of X).

Very often in applications, but not just in applications, one is interested in the probability that a process (e.g. Brownian motion) hits a barrier or exits a set; resp. one is interested in expectated values *when* these events occur. To model such events mathematically, one introduces, as is commonly known, so called *stopping times*, which are also crucial to the theory of Markov processes.

Definition 1.3.6. A random variable $\tau \colon \Omega \to [0, \infty]$ is called a *stopping time* relative to the filtration $(\mathcal{F}_t)$ iff for each t: $\{\tau \leq t\} \in F_t$.

Proposition 1.3.7. *Let $(X_t)_{t\geq 0}$ be a continuous real valued process, let (F_t^X) be its natural filtration and let $A \subset \mathbb{R}$ be closed. Define $\tau_A := \inf\{s \geq 0 \colon X_s \in A\}$. Then τ_A is a stopping time.*

Proof.

$$d(y, A) := \inf_{z\in A} |y - z| \implies x \in A \Leftrightarrow d(x, A) = 0$$

$$\tau(\omega) \leq t \Leftrightarrow \inf\{s \geq 0 \text{ s.t. } d(X_s(\omega), A) = 0\} \leq t \Leftrightarrow$$

$$\Leftrightarrow \inf_{s\leq t} d(X_s(\omega), A) = 0 \Leftrightarrow \inf_{s\leq t,\, s\in\mathbb{Q}} \underbrace{d(X_s(\omega), A)}_{h_s(\omega)} = 0 \Leftrightarrow \underbrace{\inf_{s\leq t, s\in\mathbb{Q}} h_s(\omega)}_{h(\omega)} = 0,$$

where $h_s(\omega)$ is an $\mathcal{F}_t$-measurable random variable because it is a continuous function of a r.v. This implies that $h(\omega)$, as an infimum over countably many values, is an $\mathcal{F}_t$-measurable r.v. which gives that $\{\tau \leq t\} = \{h = 0\} \in \mathcal{F}_t$. □

Remark 1.3.8. If the *usual hypotheses* are satisfied, this proposition can be generalized to Borel sets, but only at the price of high efforts in measure theory. Moreover, if we assume the filtered probability space to satisfy the *usual hypotheses* we rather elementary get:

Proposition 1.3.9. *(cf. [Pro04], p.3) Let $(\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t\geq 0})$ be a probability space with right continuous filtration. Then the event $\{\tau < t\} \in \mathcal{F}_t$, $0 \leq t \leq \infty$, if and only if τ is a stopping time.*

What we also want to consider, is the state of the process at the random time τ and therefore we define, on the set $\{\tau < \infty\}$,

$$X_\tau(\omega) := X_{\tau(\omega)}(\omega).$$

However, to ensure that $X_{\tau(\omega)}(\omega)$ really is a r.v. on $\{\tau < \infty\}$, adaptedness is not sufficient, but the process X considered as a function $T \times \Omega \to \mathbb{R}$ has to be jointly measurable in both arguments, which leads to the following

Definition 1.3.10. A real valued process X is called *progressively measurable* or *progressive* (w.r.t. the filtration $(\mathcal{F}_t)_{t\geq 0}$) if for every $t \geq 0$ the mapping

$$\Phi^t\colon [0,t] \times \Omega \to \mathbb{R}, \quad \Phi^t(s,\omega) := X_s(\omega) \tag{1.3.2}$$

is measurable w.r.t. $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$.

Proposition 1.3.11. *An adapted process with (right or left) continuous paths is progressively measurable.*

By now we know how to define the *state* of the process at a random time, but we also want to extend the concept of *information at time t* from deterministic times to *information at a random time τ*. The goal is an analogous result to adaptedness, i.e. if $(X_t)_{t\geq 0}$ is adapted, X_τ should be $\mathcal{F}_\tau$-measurable. For this reason we introduce the following σ-algebra:

Definition 1.3.12. Let $(\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t\geq 0})$ be a filtered probability space, τ a stopping time. Then $\mathcal{F}_\tau$ is the σ-algebra consisting of the sets

$$A \in \mathcal{F}_\tau \Leftrightarrow A \in \mathcal{F} \text{ and } \forall t\colon A \cap \{\tau \leq t\} \in \mathcal{F}_t. \tag{1.3.3}$$

Or equivalently: $A \in \mathcal{F}_\tau \Leftrightarrow \mathbb{1}_A = \mathbb{1}_{\{\tau \leq t\}}$ is $\mathcal{F}_t$-measurable.

Theorem 4. *Let $(X_t)_{t\geq 0}$ be a progressively measurable process and let τ be a stopping time (w.r.t. the same filtration $(\mathcal{F}_t)$) then X_τ is $\mathcal{F}_\tau$-measurable on the set $\{\tau < \infty\}$.*

Remark 1.3.13. Since for stopping times σ and τ, $(\sigma \wedge \tau)$ and $(\sigma \vee \tau)$ are stopping times; $\{\sigma = \tau\}$, $\{\sigma \leq \tau\}$, $\{\sigma < \tau\}$ are in $\mathcal{F}_\sigma \cap \mathcal{F}\tau$ and $\sigma \leq \tau$ implies $\mathcal{F}_\sigma \subseteq \mathcal{F}_\tau$, the family of σ-algebras $F_{t\wedge\tau}$ is a filtration. I.e. given a process X and a stopping time τ we can define the *stopped process* (X^τ)

$$X^\tau(\omega) := X_{t\wedge\tau}(\omega). \tag{1.3.4}$$

The notion of progressive measurability generalizes straightforward to stopping times.

Proposition 1.3.14. *If X is progressive, then X^τ is progressive w.r.t the filtration $(\mathcal{F}_{t\wedge\tau})$.*

To conclude the martingale section, we state a theorem on convergence of martingales and Doob's Optional Stopping Theorem. We skip Doob's inequalities and also some intermediary technically necessary results although the inequalities are very interesting in their own right. Crucial to the convergence of martingales is the notion of *uniform integrability* of a set of random variables, which denotes some sort of regularity that is situated between L^1 and L^p for $p > 1$.

Definition 1.3.15. A family of random variables $\{ X_i \,|\, i \in I \}$ is called *uniformly integrable (U.I.)*, if one of the two equivalent conditions holds:

(i) $\forall \varepsilon > 0 \,\exists\, M \in \mathbb{R}$ s.t $\displaystyle\int_{\{ |X_i| > M \}} |X_i| \, d\mathbb{P} \leq \varepsilon \quad \forall i \in I$

(ii) $\forall \varepsilon > 0 \,\exists\, \delta > 0$ s.t. $\forall A \in \mathcal{F}$, with $\mathbb{P}(A) \leq \delta$, $\displaystyle\int_A |X_i| \, d\mathbb{P} \leq \varepsilon$ and $\sup_{i\in I} |X_i| < \infty$.

Example 1.3.16. The trivial example consists just in the family of a single integrable r.v. X. $\{ X \}$ is U.I. But there are more interesting ones.

(i) Suppose $\exists\, Y$ s.t. $\mathbb{E}|Y| < \infty$ and $\forall i \in I$: $|X_i| \leq |Y|$. Then $\{ X_i \,|\, i \in I \}$ is U.I.
(ii) Let $p > 1$ and $\sup_{i\in I} \|X_i\|_p < \infty$. Then $\{ X_i \,|\, i \in I \}$ is U.I.
(iii) Assume that, for all $i \in I$, $\mathcal{F}_i$ is a sub-σ-algebra of $\mathcal{F}$ and that $X \in L^1$. Then $\{ \mathbb{E}[X \,|\, \mathcal{F}_i] \,|\, i \in I \}$ is U.I.

Regarding convergence, the notion of uniform integrability is particularly interesting, since it upgrades the a.s. convergence:

Theorem 5. *Let $\{ X_t \,|\, t \leq 0 \}$ be U.I. and let $\lim_{t\to\infty} X_t = X_\infty$ exist a.s. or in probability. Then $\lim_{t\to\infty} X_t = X_\infty$ in L^1.*

Theorem 6. *Let $(X_t)_{t\geq 0}$ be a cadlag martingale s.t. $\sup_{t\geq 0} \mathbb{E}[|X_t|] < \infty$. Then t.f.a.e.*

(i) $\lim_{t\to\infty} X_t$ *exists in* L^1.
(ii) There exists a r.v. X_∞ in L^1 s.t., for all $t \in [0,\infty)$, $\mathbb{E}[X_\infty \,|\, \mathcal{F}_t] = X_t$.
(iii) The family $\{X_t \,|\, t \geq 0\}$ is U.I.

Moreover, if the martingale is uniformly bounded in L^p for $p > 1$, i.e. $\sup_{t\geq 0} \|X_t\|_p < \infty$*, then all three conditions are satisfied and convergence holds in L^p.*

Theorem 7. *(Doob's Optional Stopping Theorem, cf. [RY99], p.65) Let $(X_t)_{t\geq 0}$ be a cadlag martingale.*

(i) If $\sigma \leq \tau$ are bounded stopping times, then $\mathbb{E}[X_\tau \,|\, \mathcal{F}_\sigma] = X_\sigma$.
(ii) If $\{X_t\}$ is U.I. then for every pair of not necessarily bounded stopping times $\sigma \leq \tau$ we get $\mathbb{E}[X_\tau \,|\, \mathcal{F}_\sigma] = \mathbb{E}[X_\infty \,|\, \mathcal{F}_\sigma] = X_\sigma$.

1.4 Markov Property and Processes

In Chapters 2 and 3 we will need some, but not much Markov-theory, so we introduce (again mostly without proofs) some basic notions and notations concerning Markov processes. The intuitive idea of a Markov process X, namely that at any given time s the future behaviour of $(X_t)_{t>s}$ depends only on the current state $X_s(\omega)$ will be formalized via a *transition probability* $P_{s,t}(x, A)$ that describes the probability of reaching the set A at time t, given the state $X_s = x$. Since a transition function together with an initial distribution ν of X determines the f.d.d.s of a Markov process uniquely, by Kolmogorov's extension criterion we get a canonical version of X or, in other words, we get a unique probability measure $\mathbb{P}_\nu$ on $(\mathbb{R}^{[0,\infty)}, \mathcal{B}(\mathbb{R})^{[0,\infty)})$ s.t. X is Markov w.r.t. its natural filtration and its transition function. (Again we follow closely the presentation of Revuz and Yor since it is the most concise description of that topic known to the author. We follow Revuz/Yor even in the somewhat curious integral-notation, cf. [RY99], Section III.1.)

First of all, recall the definition of the canonical version (1.1.9) and the coordinate process $Y_t(\omega) = \omega(t)$ (1.1.5). For notational convenience, we denote the state space by E, its Borel σ-algebra by $\mathcal{B}(E) = \mathcal{E}$ and $\Omega = E^{[0,\infty)}$. We define an operator on this space of functions that shifts a path to the left.

Definition 1.4.1. Let $(\Omega, \mathcal{F}) = (E^{[0,\infty)}, \mathcal{E}^{[0,\infty)})$. For $\omega \in \Omega$ the family of operators $\theta_t \colon \Omega \to \Omega$, $t \in [0, \infty)$, is defined by

$$\theta_t(\omega(s)) := \omega(t+s). \tag{1.4.1}$$

Hence, if Y is the coordinate process of a process X

$$Y_s \circ \theta_t(\omega) = Y_s(\theta_t(\omega)) = \theta_t(\omega(s)) = \omega(t+s) = Y_{t+s}(\omega). \tag{1.4.2}$$

Remark 1.4.2. The state space in this definition and for the rest of the section may be chosen to be any polish (i.e. complete, separable, metric) space E with canonical path space $\Omega = E^{[0,\infty)}$ or $\mathcal{C}([0,\infty), E)$. Moreover denote by $f \in \mathcal{E}$ ($f \in \mathcal{E}^+$ resp.) that f is a real valued $\mathcal{E}$-measurable (positive resp.) function. We always bear in mind the most intuitive setting $E = \mathbb{R}$, $\mathcal{E} = \mathcal{B}(\mathbb{R})$ and $\Omega = \mathcal{C}[0,\infty)$.

Definition 1.4.3. Let $(E, \mathcal{E})$ be a measurable space. A *transition kernel*, or *transition probability* P on E is a map $P\colon\ E \times \mathcal{E} \to [0,\infty]$, s.t.

(i) for every $x \in E$, the map $A \to P(x, A)$ is a probability measure on $\mathcal{E}$,
(ii) for every $A \in \mathcal{E}$, the map $x \to P(x, A)$ is $\mathcal{E}$-measurable.

Since $P(x, A)$ is a measure, for bounded measurable real functions f on E we can define $P(x, f) = \int_E P(x, dy) f(y)$, which is again a positive real valued measurable function on E. And for two transition kernels P and Q

$$PQ(x, A) := \int_E P(x, dy) Q(y, A)$$

is again a transition kernel.

For obvious reasons we index the transition probabilities by a time parameter $t \geq 0$. If we assume that for every $s < t$ there exists a transition kernel $P_{s,t}$ s.t.

$$\mathbb{P}(X_t \in A \,|\, \mathcal{F}_s^X) = P_{s,t}(X_s, A) \quad a.s.$$

then this is (by means of the standard machinery, i.e. taking simple functions and considering limits) equivalent to the statement, that for all $f \in \mathcal{E}^+$

$$\mathbb{E}[f(X_t) \,|\mathcal{F}_s^X] = P_{s,t}(X_s, f).$$

However, in order to determine how a process evolves along the positive real line $[0,\infty)$, it's necessary to know for every three points $s < v < t$ the probabilities of getting from the state at time s to the state at time t, given the transition probabilities of getting from s to v and from v to t. These heuristics are captured by the *Chapman-Kolmogorov* equation and the following definition which in addition enables us to define a Markov process (cf. [RY99], p. 75f.).

Definition 1.4.4. A *transition function* (t.f.) on $(E, \mathcal{E})$ is a family $P_{s,t}$, $0 \le s < t$ of transition probabilities in $(E, \mathcal{E})$ such that for every triple $s < v < t$ in $\mathbb{R}^+$ we have

$$P_{s,t}(x, A) = \int_E P_{s,v}(x, dy)\, P_{v,t}(y, A) \tag{1.4.3}$$

for all $x \in E$ and $A \in \mathcal{E}$. A t.f. is said to be *homogeneous*, if $P_{s,t}$ depends on t and s only through the difference $t - s$. By convention then $P_{0,t} = P_t$ and the Chapman-Kolmogorov equation reads

$$P_{s+t}(x, A) = \int_E P_s(x, dy)\, P_t(y, A) \tag{1.4.4}$$

for every $s, t \ge 0$. I.e., $\{P_t;\, t \ge 0\}$ is a (stochastic) semi-group.

Definition 1.4.5. Let $(\Omega, \mathcal{F}, (\mathcal{G}_t), \mathbb{P})$ be a filtered probability space. An adapted process X is a *Markov process* w.r.t. $(\mathcal{G}_t)$ and with t.f. $P_{s,t}$ if for any $f \in \mathcal{E}^+$ and for any pair (s, t), $s < t$

$$\mathbb{E}[f(X_t) \,|\, \mathcal{G}_s] = P_{s,t}(X_s, f). \quad a.s.$$

The probability measure $\nu = X_0(\mathbb{P})$ is called the *initial distribution* of X. The process is said to be homogeneous if the t.f. is homogeneous and the above equation simplifies to

$$\mathbb{E}[f(X_t) \,|\, \mathcal{G}_s] = P_{t-s}(X_s, f).$$

Remark 1.4.6. The canonical example for a t.f. is, as usual, given by Brownian motion. Let $E = \mathbb{R}$, $\mathcal{E} = \mathcal{B}(\mathbb{R})$ and consider $P_t(x, \cdot)$ as the measure with Lebesgue-density

$$g_{t,x}(y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x)^2}{2t}}.$$

Then this family of kernels is, as can be shown directly by checking the Chapman-Kolmogorov equation, a homogeneous transition function.

However, it's not so easy to show directly that a process defined via a t.f. satisfies the above definition of a Markov process. To ensure that, a somewhat more tractable characterization is needed. It should also be noted that if X is Markov w.r.t. some filtration $(\mathcal{G}_t)$ it's also Markov w.r.t. to its natural filtration $(\mathcal{F}_t^X)$.

The following proposition is the key result to construct a canonical version of a Markov process, given a transition function.

Proposition 1.4.7. *(cf. [RY99], p.76) A process X is Markov w.r.t. to its natural filtration $(\mathcal{F}_t^X)$ with t.f. $P_{s,t}$ and initial measure ν if and only if for any $0 = t_0 < t_1 \cdots < t_k$ and for all $A_0, \dots, A_k \in \mathcal{E}$ we have*

$$\mathbb{P}(X_{t_0} \in A_0, \dots, X_{t_k} \in A_k) = \int_{A_0} \nu(dx_0) \underbrace{\int_{A_1} P_{t_0,t_1}(x_0, dx_1)}_{P_{t_0,t_1}(x_0,A_1)} \cdots \int_{A_k} P_{t_{k-1},t_k}(x_{k-1}, dx_k). \tag{1.4.5}$$

Theorem 8. *Given a transition function $P_{s,t}$ on $(E, \mathcal{E})$; for any probability measure ν on $(E, \mathcal{E})$ there is a unique probability measure $\mathbb{P}^\nu$ on $(E^{[0,\infty)}, \mathcal{E}^{[0,\infty)})$ s.t. X is Markov w.r.t. its natural filtration, with t.f. $P_{s,t}$ and initial measure ν.*

Proof. Define the family of the f.d.d.s by

$$\mathcal{P}_X^f := \{\mathbb{P}^\nu_{t_1,\dots,t_k} \,|\, k \ \in \mathbb{N},\, t_i \in \mathbb{R}^+\},$$

where $\mathbb{P}^\nu_{t_1,\dots,t_k}(A_1 \times \cdots \times A_k) = \mathbb{P}(X_{t_0} \in A_0, \dots, X_{t_k} \in A_k)$ is defined by (1.4.5) above. By Kolmogorov's extension criterion and Proposition 1.4.7 X is Markov for the probability measure $\mathbb{P}^\nu$. □

Remark 1.4.8. From this construction it is clear that Brownian motion started at an arbitrary point $x \in \mathbb{R}$ is a Markov process and that we can identify $\mathbb{P}^x$ defined in Remark 1.2.4 and $\mathbb{P}^{\delta_x}$, using $\mathbb{P}^x$ as an abbreviation for the case $\nu = \delta_x(A)$, where δ_x is the Dirac measure. The expectation w.r.t. $\mathbb{P}^x$ ($\mathbb{P}^\nu$ resp.) is denoted by $\mathbb{E}^x$ ($\mathbb{E}^\nu$ resp.). Moreover, for a homogeneous t.f. we get $\mathbb{P}^x(X_t \in A) = P_t(x, A)$ and by the definition of a transition kernel, the map $x \mapsto \mathbb{P}^x(X_t \in A)$ is measurable. This is needed in particular to compute the expectation w.r.t. the measure $\mathbb{P}^\nu$.

Proposition 1.4.9. *Let Y be $\mathcal{F}_\infty^X$ measurable and positive or bounded, then the map $x \mapsto \mathbb{E}^x[Y]$ is $\mathcal{E}$-measurable and*

$$\mathbb{E}^\nu[Y] = \int_E \nu(dx)\, \mathbb{E}^x[Y]. \tag{1.4.6}$$

Now we are ready to state one main result of this section which consists in a convenient formulation of the Markov property.

Proposition 1.4.10. *(Markov property. cf. [RY99], p.78) Let Y be $\mathcal{F}_\infty^X$-measurable and positive (or bounded). Then, for every $t > 0$ and any initial measure ν,*

$$\mathbb{E}^\nu[Y \circ \theta_t \,|\, \mathcal{F}_t^X] = \mathbb{E}^{X_t}[Y]. \quad \mathbb{P}^\nu \ a.s. \tag{1.4.7}$$

Remark 1.4.11. To take full advantage of the above proposition set $Y = f(X_t)$ for any bounded measurable f. Then, if the (coordinate) process X has the Markov property, we obtain

$$\mathbb{E}[f(X_{s+t}) \,|\, \mathcal{F}_s] = \mathbb{E}^{\nu}[f(X_t) \circ \theta_s \,|\, \mathcal{F}_s] = \mathbb{E}^{X_s}[f(X_t)]. \tag{1.4.8}$$

And by the definition of a Markov process

$$\mathbb{E}[f(X_{s+t}) \,|\, \mathcal{F}_s] = P_t(X_s, f) = \mathbb{E}^{X_s}[f(X_t)]. \tag{1.4.9}$$

It should also be noted that so far we only worked with the natural filtrations. In view of Remark 1.3.8, however, it is necessary to consider filtrations satisfying the usual hypotheses, i.e. filtrations being complete and right continuous. Here we only note, that all results above (w.r.t. the natural filtrations) hold also if we complete the natural filtrations w.r.t. all initial distributions ν, thereby making the filtration already right continuous.

In order to obtain some further results like the strong Markov property and to obtain more tractable versions of a Markov process we will have to put stronger assumptions on the transition function P_t, regarded as an operator on the space $\mathcal{C}_0(E)$ of continuous functions vanishing at infinity. (The state space E is assumed to be locally compact with countable base.) Therefore we will define Feller processes and just note some important results without developing the theory.

Definition 1.4.12. A Feller semigroup on $C_0(E)$ is a family T_t, $t \geq 0$, of positive linear operators on $C_0(E)$ such that

(i) $T_0 = Id$ and $\|T_t\| \leq 1$ for every t.
(ii) $T_{t+s} = T_t \circ T_s$ for any pair $s, t \geq 0$.
(iii) $\lim_{t \searrow 0} \|T_t f - f\| = 0$ for every $f \in C_0(E)$.

Proposition 1.4.13. *(cf. [RY99], p.83)*
With each Feller semigroup on E one can associate a unique homogeneous transition function P_t, $t \geq 0$, on $(E, \mathcal{E})$ such that

$$T_t f(x) = P_t f(x) = P_t(x, f)$$

for every $f \in C_0(E)$ and every $x \in E$.

Crucial to the theory of Feller processes is the following

Theorem 9. *(cf. [RY99], p.86)*
A Feller process $(X_t)_{t\geq 0}$ admits a cadlag modification.

Another important feature of Feller processes is the strong Markov property. In order to define the strong Markov property we have to extend the definition of the shift operators $\theta_t\colon \Omega \to \Omega$ to stopping times T. We set $\theta_T(\omega) = \theta_t(\omega)$ if $T(\omega) = t$, s.t.

$$X_t \circ \theta_T = X_{T+t} \text{ and } \theta_T^{-1}(\mathcal{F}_\infty^X) \subset \sigma(X_{T+t}, t \geq 0).$$

Theorem 10. *(Strong Markov property)*
Let (X_t) be the canonical cadlag version of a Feller process, let Z be a $\mathcal{F}_\infty$-measurable and nonnegative (or bounded) random variable and let T be a finite $\mathcal{F}_t$-stopping time, then for any initial measure ν

$$\mathbb{E}^\nu[Z \circ \theta_T \,|\, \mathcal{F}_T] = \mathbb{E}^{X_T}[Z]. \quad \mathbb{P}^\nu \; a.s.$$

Remark 1.4.14. The strong Markov property can also be defined without reference to a transition function. We will need the following definition in Chapter 3.

Definition 1.4.15. (Cf. [Low08a].) A real valued process X is strong Markov if for every bounded measurable $g\colon \mathbb{R} \to \mathbb{R}$ and every $t > 0$ there exists a measurable $f\colon \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ s.t. for every finite stopping time T

$$f(T, X_T) = \mathbb{E}[g(X_{T+t}) \,|\, \mathcal{F}_T].$$

Remark 1.4.16. The uniform continuity property in the definition of a Feller semigroup is actually equivalent to pointwise continuity which can, for instance, be shown easily for the t.f. of Brownian motion; i.e. BM is a Feller process and hence strong Markov.

1.5 Half a nutshell of Lévy processes

Brownian motion, as we saw, enjoys the (strong) Markov property which is not a big surprise since it has stationary and independent increments. Conversely, if we take the property of having stationary and independent increments as a definition, we get a very large (and venerable) class of stochastic processes, namely Lévy processes. (Cf. [RY99], p.91f and p.109f for most parts of the following.)

Definition 1.5.1. (Cf. [Pro04], p.20) Let $(X_t)_{t\geq 0}$ be a real valued adapted process, $X_0 = 0$. X is called a *Lévy process* if

(i) X has stationary independent increments, and
(ii) X is continuous in probability, i.e. $\lim_{t\to s} \mathbb{P}(|X_t - X_s| > \varepsilon) = 0$ for every $\varepsilon > 0$.

With regard to Section 2.3, where we will construct a mimicking process via a transition function, we take a look on an interesting class of families of probability measures, namely convolution semigroups. They provide a link between Feller processes and Lévy processes.

Definition 1.5.2. Let $(\mu_t)_{t\geq 0}$ be a family of probability measures on $\mathbb{R}^d$. Then $(\mu_t)_{t\geq 0}$ is called a *convolution semigroup* iff

(i) $\mu_t * \mu_s = \mu_{t+s}$ for any pair (s,t),
(ii) $\mu_0 = \delta_0$ and $\lim_{t\searrow 0} \mu_t = \delta_0$ in the vague topology.

Lemma 1.5.3. *Let $(\mu_t)_{t\geq 0}$ be a convolution semigroup and define*

$$P_t(x, A) := \int_{\mathbb{R}^d} \mathbb{1}_A(x+y)\mu_t(dy).$$

Then (P_t) is a Feller transition function.

Proposition 1.5.4. *If the t.f. of a Feller process X is given by a convolution semigroup (μ_t) then X has stationary independent increments, i.e. is a Lévy process. The law of the increment $X_t - X_s$ is μ_{t-s}.*

Proof. For any $f \in \mathcal{E}^+$ and any t we have, since $\mathbb{P}^x(X_0 = x) = 1$,

$$\mathbb{E}^x[f(X_t - X_0)] = \mathbb{E}^x[f(X_t - x)] = \mu_t(f),$$

which no longer depends on x. By the Markov property we get for $s < t$

$$\mathbb{E}^\nu[f(X_t - X_s) \,|\, \mathcal{F}_s] = \mathbb{E}^{X_s}[f(X_{t-s} - X_0)] = \mu_{t-s}(f) \quad \mathbb{P}^\nu \text{ a.s.}$$

$\square$

Remark 1.5.5. If a Feller process has stationary independent increments, then it can also be proved that its transition function is given by a convolution semigroup.

Another important tool is the notion of infinite divisibility of a random variable X, or a probability measure μ on $\mathbb{R}$, resp. $\mathbb{R}^d$.

Definition 1.5.6. A real (or $\mathbb{R}^d$) valued random variable X is said to be *infinitely divisible* if for every $n \in \mathbb{N}$ there exist n independent identically distributed random variables $Y_i^{(n)}$ such that $X \overset{(law)}{=} Y_1^{(n)} + \cdots + Y_n^{(n)}$.
Likewise, a probability measure μ is said to be infinitely divisible if there exists a probability measure μ_n such that $\mu = \mu_n^{*n}$, i.e. μ is the n-fold convolution of μ_n.

Proposition 1.5.7. *Let $(X_t)_{t\geq 0}$ be a Lévy process, then X_t is infinitely divisible for each $t \geq 0$.*

Proof. For all $n \in \mathbb{N}$ we can write X_t as

$$X_t = (X_t - X_{(1-\frac{1}{n})t}) + (X_{(1-\frac{1}{n})t} - X_{(1-\frac{2}{n})t}) + \cdots + (X_{\frac{1}{n}t} - X_0).$$

By definition of a Lévy process these increment are iid and so X_t is infinitely divisible. $\square$

Remark 1.5.8. Conversely it can be shown that any infinitely divisible r.v. Y may be imbedded in a unique convolution semigroup, i.e., there is a Lévy process X s.t. $Y \overset{(law)}{=} X_1$. Therefore we need the famous and fundamental *Lévy Khintchine formula* (here stated just for dimension one):

Theorem 11. *(Cf. [Sat99], p.37; [RY99], p.110) A probability measure μ on $\mathbb{R}$ is infinitely divisible if and only if its Fourier transform $\hat{\mu}$ is equal to $\exp(\psi)$ with*

$$\psi(u) = i\beta u - \frac{\sigma^2 u^2}{2} + \int_{\mathbb{R}\setminus\{0\}} \left(e^{iux} - 1 - \frac{iux}{1+x^2}\right) \nu(dx), \tag{1.5.1}$$

where $\beta \in \mathbb{R}$, $\sigma \geq 0$ and ν is a Radon measure on $\mathbb{R}\setminus\{0\}$ such that

$$\int \frac{x^2}{1+x^2}\, \nu(dx) < \infty.$$

Remark 1.5.9. The measure ν in the characteristic exponent ψ is known as the Lévy measure of X and it accounts for the jumps of X. However, to define a Lévy process X based on an infinitely divisible measure, observe that the above formula amounts

to say that, for each $t \geq 0$, $\exp(t\,\psi(u))$ is the Fourier transform of a probability measure μ_t, or equivalently, $\exp(t\,\psi(u)) = \mathbb{E}[\exp(iuX_t)] = \varphi_{X_t}(u)$ is the characteristic function of a r.v. $X_t \sim \mu_t$. Clearly $\mu_t * \mu_s = \mu_{t+s}$ and $\lim_{t\searrow 0}\mu_t = \delta_0$, hence (μ_t) is a convolution semigroup and therefore defines a Lévy process according to Proposition 1.5.4.
Note that in general the characteristic triplet (β, σ, ν) determines uniquely an infinitely divisible measure μ, or a Lévy process respectively. The parameter β accounts for the drift of the process, σ for the (Brownian) diffusion term and the Lévy measure ν for the jumps. Hence a Lévy process is continuous if and only if $\nu \equiv 0$. If $X_t = bt + \delta B_t$ we know from the normal distribution that $\psi(u) = iua - \frac{1}{2}\sigma^2 u^2$, hence $\beta = b$ and $\sigma = \delta$.

In Section 2.3 a particular class of Lévy processes will be particularly important for us, namely the *subordinators* which are a.s. increasing Lévy processes and the law of which is thus carried by $[0,\infty)$. Hence for subordinators not only the characteristic but also the moment generating function is defined and one gets a special case of the Lévy Khintchine formula. (The canonical example of a subordinator is the Poisson process.)

Definition 1.5.10. (cf. [KS88], p.405) A real valued process $N = (N_t)_{t\in[0,\infty)}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *subordinator* if it has stationary independent increments and almost every path of N is non-decreasing, right continuous and satisfies $N_0 = 0$.

Theorem 12. *(Lévy, Khintchine, Itō; cf. [KS88], p.405) The moment generating function of a subordinator $N = (N_t)_{t\in[0,\infty)}$ is given by $\mathbb{E}[e^{-uN_t}] = \exp(-t\,\phi(u))$, where*

$$\phi(u) = \beta u + \int_{(0,\infty)} (1 - e^{-ux})\,\nu(dx), \tag{1.5.2}$$

$u \geq 0$, β a constant in $\mathbb{R}^+$ and ν is a σ-finite measure on $(0,\infty)$ for which the above integral is finite.

Remark 1.5.11. According to the foregoing remark, the characteristic triplet in the case of subordinators reads $(\beta, 0, \nu)$. The Brownian part σ has to be zero since the process is increasing almost surely. Moreover the role of the characteristic exponent ψ here is taken by what is called the Laplace exponent ϕ.
Having the above in mind, it is remarkable that a Lévy process (in the sense of its f.d.d.s) is fully determined by its characteristic function, or equivalently, that the law of a Lévy process is determined solely by its one dimensional marginals.

1.6 Stochastic integration

Roughly, there are three approaches to stochastic integration, i.e three ways to define a stochastic integral. In the order of increasing generality they read as follows: first, as done by K. Itō, one can define the stochastic integral w.r.t. Brownian motion; then, as done by Kunita-Watanabe 1967, for general square integrable martingales (cf. e.g. [KS88]) and in the most general case it is also possible to define the stochastic integral w.r.t. (not necessarily continuous) semimartingales. (Compare e.g. [Pro04]).

We won't need the general case of discontinuous semimartingales, so we restrict ourselves to the classical Itō-integral and stochastic integration w.r.t. continuous local martingales. Furthermore we will not develop the theory of stochastic integration but just sketch a possible construction and state the definitions, properties and important results. Compare [RY99] (Chapter IV), [KS88] (Chapter 3) or [Pro04] (Chapter II) for three different approaches. We shall mainly follow [Øks98] and [RY99].

1.6.1 Preliminaries & Quadratic (Co)-Variation

Definition 1.6.1. Let $[0,t]$ be a compact interval and let $\Delta = \{0 = t_0 < t_1 \cdots < t_m = t\}$ be a subdivision of $[0,t]$. Let $(X_t)_{t\geq 0}$ be an adapted stochastic process and define

$$T_t^{\Delta}(X) = \sum_{t_j\in\Delta} (X_{t_{j+1}} - X_{t_j})^2.$$

The process X is said to be of finite quadratic variation if, for all $t \geq 0$ and all subdivisions Δ^n of $[0,t]$ such that $|\Delta^n| \to 0$, the limit

$$\langle X_t, X_t\rangle = \mathbb{P} - \lim_{n\to\infty} T_t^{\Delta^n}(X) \tag{1.6.1}$$

exists. We will denote the quadratic variation process $\langle X, X\rangle_t$ of X also by $\langle X\rangle_t$.

Proposition 1.6.2. *Let $(B_t)_{t\geq 0}$ be a standard Brownian motion, then $\langle B, B\rangle_t = t$.*

Proof. We will show convergence in L^2 which includes convergence in probability. Let $t \geq 0$ and let Δ^n be a subdivision of $[0,t]$ s.t. $|\Delta^n| \to 0$. Recall that the increments of Brownian motion are independent, hence in $L^2(\Omega, \mathbb{P})$ we have

$$\left((B_{t_{i+1}} - B_{t_i})^2 - (t_{i+1} - t_i)\right) \perp \left((B_{t_{j+1}} - B_{t_j})^2 - (t_{j+1} - t_j)\right)$$

for $i \neq j$ and $t_i, t_j \in \Delta$, i.e., we can apply Pythagoras. Recall also the distribution of BM, $(B_t - B_s) \sim \mathcal{N}(0, t-s)$. In particular we get:

$$\begin{aligned}
\left\|T_t^{\Delta^n}(B) - t\right\|_2^2 &= \left\|\sum_{t_j \in \Delta^n} (B_{t_{j+1}} - B_{t_j})^2 - t\right\|_2^2 = \left\|\sum_{t_j \in \Delta^n} (B_{t_{j+1}} - B_{t_j})^2 - (t_{j+1} - t_j)\right\|_2^2 \\
&= \sum_{t_j \in \Delta^n} \left\|(B_{t_{j+1}} - B_{t_j})^2 - (t_{j+1} - t_j)\right\|_2^2 = \\
&= \sum_{t_j \in \Delta^n} \mathbb{E}\left[\left((B_{t_{j+1}} - B_{t_j})^2 - (t_{j+1} - t_j)\right)^2\right] = \\
&= \sum_{t_j \in \Delta^n} 2(t_{j+1} - t_j)^2 \leq 2 \sup_{t_j \in \Delta^n} |t_{j+1} - t_j| \sum_{t_j \in \Delta^n} |t_{j+1} - t_j| = 2\Delta^n t.
\end{aligned}$$

Hence we get that $\left\|T_t^{\Delta^n}(B) - t\right\|_2^2 \to 0$ for $\Delta^n \to 0$. □

Remark 1.6.3. As commonly known, precisely this non-zero quadratic variation is the reason why the stochastic integral w.r.t. Brownian motion cannot be defined simply as a pathwise Lebesgue-Stieltjes integral.

Definition 1.6.4. Let X and Y be two stochastic processes, let Δ be a subdivision of $[0, t]$ then the *cross variation* or *quadratic covariation* process $\langle X, Y \rangle_t$ is defined as limit in probability

$$\langle X, Y \rangle_t = \lim_{|\Delta| \to 0} \sum_i (X_{t_{i+1}} - X_{t_i})(Y_{t_{i+1}} - Y_{t_i}).$$

The following theorem gives another hint why quadratic variation plays a crucial role in stochastic calculus.

Theorem 13. *(cf. [RY99], p.120) A continuous and bounded martingale M is of finite quadratic variation and $\langle M, M \rangle_t$ is the unique continuous increasing adapted process vanishing at zero, s.t. $M^2 - \langle M, M \rangle$ is a martingale.*

Definition 1.6.5. An adapted stochastic process $(M_t)_{t \geq 0}$ is a (continuous) *local martingale* if there exists an increasing sequence of stopping times $(T_n)_{n \geq 1}$, $T_n \uparrow +\infty$, such that $X_{T_n \wedge t}$ is a (continuous) martingale for all n.

Definition 1.6.6. A process $(X_t)_{t\geq 0}$ is a (continuous) *semimartingale* if it is the sum of a (continuous) local martingale $(M_t)_{t\geq 0}$ and a (continuous) process $(A_t)_{t\geq 0}$ of locally bounded finite variation (i.e. $\mathrm{Var}_t(A) < \infty$ for all t, where $\mathrm{Var}_t(A)$ is the variation of A on $[0,t]$).

Remark 1.6.7. The decomposition of a semimartingale into a local martingale and an adapted process of finite variation is unique as long as one does not change the filtration. (Otherwise there may, of course, be different decompositions.)

Proposition 1.6.8. *(cf. [RY99], p.128) A continuous semimartingale $X = M + A$ is of finite quadratic variation and $\langle X, X\rangle_t = \langle M, M\rangle_t$.*

Proof. We prove that the cross variation $\langle M, A\rangle_t$ is zero as well as the QV of A. Let Δ be a subdivision of $[0,t]$, then

$$\left|\sum_i (M_{t_{i+1}} - M_{t_i})(A_{t_{i+1}} - A_{t_i})\right| \leq \left(\sup_i M_{t_{i+1}} - M_{t_i}\right) \mathrm{Var}_t(A),$$

and this converges to zero if $|\Delta| \to 0$, because M is continuous. Likewise

$$\lim_{|\Delta|\to 0} \left(\sum_i (A_{t_{i+1}} - A_{t_i})^2\right) = 0.$$

□

Theorem 14. *(cf. [RY99], p.124) If M is a continuous local martingale, there exists a unique increasing continuous process $\langle M, M\rangle$ vanishing at zero, such that $M^2 - \langle M, M\rangle$ is a continuous local martingale and this process coincides with the quadratic variation of M in the sense of a limit in probability.*

Remark 1.6.9. In general, the *bracket* or *increasing process* of M does not coincide with quadratic variation. In the case of continuous local martinales, however, even the quadratic covariation coincides with the (co)bracket process. Furthermore, since the bracket process is continuous and increasing, it is perfectly clear, that one can define the Lebesgue-Stieltjes integral w.r.t. to $d\langle M, M\rangle_s$ or $d\langle M, N\rangle_s$. And as we will see in the Itō-formula, it is precisely an additional integral-term w.r.t. the bracket process that makes the difference between ordinary and stochastic calculus.

Theorem 15. *(cf. [RY99], p.125) If M and N are two continuous local martinales, there exists a unique continuous process $\langle M, N, \rangle$ of finite variation vanishing at zero,* *such that $MN - \langle M, N\rangle$ is a continuous local martingale. The process $\langle M, N\rangle$ coincides with the cross-variation process in probability. Furthermore, the polarization identity holds:*

$$\langle M, N\rangle = \frac{1}{4}\Big(\langle M+N, M+N\rangle - \langle M-N, M-N\rangle\Big).$$

1.6.2 The Itō-integral

We briefly sketch the easiest way to construct the Itō-integral (in one dimension), namely as a limit in $L^2([0,T]\times\Omega)$.

Let B be a linear Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t\geq 0})$. Let c be an adapted left continuous simple process, i.e. a piecewise constant function $c\colon [0,T]\times\Omega \to \mathbb{R}$ such that

$$c(t,\omega) = \sum_{i=0}^{n-1} c_i(\omega)\mathbb{1}_{[t_i,t_{i+1})},$$

where $0 = t_0 \leq t_1 \leq \cdots \leq t_n = T$ and c_i is $\mathcal{F}_{t_i}$-measurable. Then define the Itō-integral of the simple process $c(t,\omega)$ as

$$I(c)(\omega) = \int_0^T c(t,\omega)\, dB_t(\omega) = \sum_{i=0}^{n-1} c_i(\omega)\left(B_{t_{i+1}}(\omega) - B_{t_i}(\omega)\right). \tag{1.6.2}$$

We observe that, for bounded c, we get the following isometry between $L^2([0,T]\times\Omega)$ and $L^2(\Omega)$:

Lemma 1.6.10. *(Itō-isometry, cf. [Øks98], p.26) Let the simple process $c(t,\omega)$ be bounded, then*

$$\mathbb{E}\left[\left(\int_0^T c(t,\omega)\, dB_t(\omega)\right)^2\right] = \mathbb{E}\left[\int_0^T c(t,\omega)^2\, dt\right]. \tag{1.6.3}$$

This isometry now is used to extend the definition of the integral to all adapted f satisfying

$$\mathbb{E}\left[\left(\int_0^T f(t,\omega)\, dB_t(\omega)\right)^2\right] < \infty, \tag{1.6.4}$$

via a limit in $L^2(\Omega)$.

Definition 1.6.11. (Itō-integral, cf. [Øks98], p.29) Let f satisfy the above condition, then the Itō-integral of f from 0 to T is defined by

$$I(f)(\omega) = \int_0^T f(t,\omega)\, dB_t(\omega) = \lim_{n\to\infty} \int_0^T \phi_n(t,\omega)\, dB_t(\omega),$$

where the limit is taken in $L^2(\Omega)$ and where (ϕ_n) is a sequence of simple processes s.t.

$$\mathbb{E}\left[\int_0^T (f(t,\omega) - \phi_n(t,\omega))^2\, dt\right] \to 0 \quad \text{for } n \to \infty.$$

Proposition 1.6.12. *(Basic properties of the Itō-integral) Let f and g satisfy (1.6.4), let $0 \leq s < u < t$, let $\lambda \in \mathbb{R}$ then:*

(i) $\int_s^t f\, dB_v = \int_s^u f\, dB_v + \int_u^t f\, dB_v$.
(ii) $\int_s^t (\lambda f + g)\, dB_v = \lambda \int_s^t f\, dB_v + \int_s^t g\, dB_v$ *for a.a.* ω.
(iii) $\mathbb{E}[\int_s^t f\, dB_v] = 0$.
(iv) $\int_s^t f\, dB_v$ *is* $\mathcal{F}_t$*-measurable.*
(v) $\mathbb{E}[\int_0^t f\, dB_v \,|\, \mathcal{F}_s] = \int_0^s f\, dB_v$.

Proof. All of the above properties hold almost trivially in the case of simple processes; to get the martingale property (v), one has to use elementary features of conditional expectation. Because of the Itō-isometry the above properties can, by taking limits, be extended to all f, g satisfying (1.6.4). □

1.6.3 Itō-formula & Itō-processes

We start with the one-dimensional formula w.r.t. Itō-processes then state the multi-dimensional Itō-formula and present the Martingale Representation Theorem in one dimension.

Definition 1.6.13. (Cf. [Øks98], p.43.) Let $(B_t)_{t\geq 0}$ be a standard Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$. A one-dimensional Itō-process is a stochastic process of the form

$$X_t = X_0 + \int_0^t \sigma(s,\omega)\, dB_s + \int_0^t b(s,\omega)\, ds, \tag{1.6.5}$$

where σ and b are adapted and $\mathbb{P}[\int_0^t \sigma(s,\omega)^2 + |b(s,\omega)|\, ds < \infty] = 1$ for all $t \geq 0$. Often we will abbreviate the above decomposition by using the differential notation

$$dX_t = \sigma(t,\omega)\, dB_t + b(t,\omega)\, dt, \tag{1.6.6}$$

but keep in mind, that we are talking of integrals.

The Itō-integral constructed in the previous section can now be extended to integrals w.r.t. to Itō-processes.

Definition 1.6.14. Let X be an Itō-process of the form (1.6.5), let (H_s) be an adapted process, such that

$$\mathbb{P}[\int_0^t H(s,\omega)^2\sigma(s,\omega)^2 + |H(s,\omega)b(s,\omega)|\,ds < \infty] = 1, \quad \forall t \geq 0,$$

then the integral of H w.r.t. X is defined as

$$\int_0^t H(s,\omega)\,dX_s = \int_0^t H(s,\omega)\sigma(s,\omega)\,dB_s + \int_0^t H(s,\omega)b(s,\omega)\,ds. \tag{1.6.7}$$

Proposition 1.6.15. *(Itō-formula, cf. [Øks98], p.44) Let (X_t) be a one dimensional Itō-process, $dX_t = \sigma(t,\omega)\,dB_t + b(t,\omega)\,dt$, and let $f(t,x) \in \mathcal{C}^{1,2}([0,\infty) \times \mathbb{R})$, then $Y_t = f(t, X_t)$ is again an Itō-process and*

$$dY_t = \frac{\partial}{\partial t} f(t,X_t)\,dt + \frac{\partial}{\partial x} f(t,X_t)\,dX_t + \frac{1}{2}\frac{\partial^2}{\partial x^2} f(t,X_t)\,d\langle X,X\rangle_t. \tag{1.6.8}$$

Remark 1.6.16. Note that (Itō)-processes of the form (1.6.5) are semimartingales with $M_t = X_0 + \int_0^t \sigma(s,\omega)\,dB_s$ and $A_t = \int_0^t b(s,\omega)\,ds$. Hence via Proposition 1.6.8 and the Itō-isometry we obtain

$$\langle X,X\rangle_t = \langle M,M\rangle_t = \langle \int \sigma(s,\omega)\,dB_s, \int \sigma(s,\omega)\,dB_s\rangle_t = \int_0^t \sigma(s,\omega)^2\,ds,$$

or in differential form,

$$d\langle X,X\rangle_t = \sigma(t,\omega)^2\,dt.$$

We now extend the definition of an Itō-process to dimension n.

Definition 1.6.17. Let $(B_t) = (B_t^1, \dots, B_t^m)$ be an m-dimensional Brownian motion and let $u_i(t,\omega)$ and $v_{ij}(t,\omega)$, $i \in \{1,\dots,n\}$, $j \in \{1,\dots,m\}$, be adapted processes which satisfy the integrability conditions for one-dimensional Itō-processes. Then the process X defined by

$$dX_t = \begin{cases} dX_t^1 & = u_1\,dt & + & v_{11}\,dB_t^1 + \dots + v_{1m}\,dB_t^m \\ \vdots & \vdots & \vdots & \vdots \\ dX_t^n & = u_n\,dt & + & v_{n1}\,dB_t^1 + \dots + v_{nm}\,dB_t^m \end{cases}$$

is called an *m-dimensional Itō-process*. The obvious matrix notation reads

$$dX_t = u(t,\omega)\,dt + v(t,\omega)\,dB_t.$$

Theorem 16. *(Multidimensional Itō-formula, cf. [Øks98], p.48) Let*

$$dX_t = u(t,\omega)\,dt + v(t\omega)\,dB_t$$

be an n-dimensional Itō-process and let $f(t,x) = (f_1(t,x),\dots,f_d(t,x))$ be continuously differentiable, $f \in \mathcal{C}^{1,2}([0,\infty)\times\mathbb{R}^n,\mathbb{R}^d)$. Then the process

$$Y_t(\omega) = f(t,X_t)$$

is again an Itō-process and the k-th entry in the vector $Y_t(\omega) = (Y_t^1(\omega),\dots,Y_t^d)$ is given by:

$$dY_t^k = \frac{\partial}{\partial t} f_k(t,X_t)\,dt + \sum_i \frac{\partial}{\partial x_i} f_k(t,X_t)\,dX_t^i + \frac{1}{2}\sum_{i,j}\frac{\partial^2}{\partial x_i \partial x_j} f_k(t,X_t)\,d\langle X^i, X^j\rangle_t.$$

Remark 1.6.18. The (multidimensional) Itō-formula is also valid when stated for (general, e.g. discontinuous) semimartingales instead of Itō-processes. (Cf. [RY99], p.146ff. for the case of continuous semimartingales, and [Pro04], p.78ff. for the general case.)

Before we turn to the next section devoted to SDEs, we state the converse result to the fact that every Itō-integral is a martingale, namely that every martingale adapted to the Brownian filtration can uniquely be written as an Itō-integral. This holds for the n-dimensional case, although we state the theorem just in dimension one.

Theorem 17. *(Martingale Representation Theorem, cf. [Øks98], p.53) Let $(B_t)_{t\geq 0}$ be a standard Brownian motion on $(\Omega,\mathcal{F},\mathbb{P};(\mathcal{F}_t)_{t\geq 0})$, where $\mathcal{F}_t = \sigma(B_s; s\leq t)$, and let $(M_t)_{t\geq 0}$ be an $(\mathcal{F}_t)$-martingale, $M_t \in L^2(\Omega)$ for all $t\geq 0$. Then there exists a unique adapted process $\Gamma(s,\omega)$, $\mathbb{E}[\int_0^t \Gamma(s,\omega)^2\,ds] < \infty$ for all $t \geq 0$, such that:*

$$M_t(\omega) = \mathbb{E}[M_0] + \int_0^t \Gamma(s,\omega)\,ds \qquad \textit{a.s. and for all } t \geq 0. \tag{1.6.9}$$

1.7 SDEs – Generators

To begin with, we note that we will only discuss *diffusion processes*, i.e. 'Markov processes which have continuous sample paths and can be characterized in terms of its infintesimal generator', how Karatzas/Shreve denote them loosely speaking (cf. [KS88], p.281).

First, we state the (pathwise) existence and uniqueness result of Itō, in the second subsection we briefly take a look on solutions given via the so called martingale problem proposed by Stroock and Varadhan. In this section we will mainly follow [Bas98] and [Øks98], although we consider almost all results in the time-inhomogeneous case, whereas Bass and Øksendal mainly are concerned with homogeneous SDEs.

1.7.1 Pathwise solutions

Let $(B_t)_{t\geq 0}$ be a linear Brownian motion and let $\sigma\colon [0,\infty)\times\mathbb{R}\to\mathbb{R}$ and $b\colon [0,\infty)\times\mathbb{R}\to\mathbb{R}$ be measurable functions, then we will be concerned with the *stochastic differential equation* (SDE)

$$dX_t = \sigma(t,X_t)\,dB_t + b(t,X_t)\,dt, \quad X_0 = x. \tag{1.7.1}$$

As we know from Itō-calculus above, this is just an abbreviation for the integral equation

$$X_t = x + \int_0^t \sigma(s,X_s)\,dB_s + \int_0^t b(s,X_s)\,ds. \tag{1.7.2}$$

In the higher dimensional case we only consider equations with quadratic diffusion matrix. Let $(B_t)_{t\geq 0}$ be an n-dimensional Brownian motion, let $\sigma\colon [0,\infty)\times\mathbb{R}^n\to\mathbb{R}^{n\times n}$ and $b\colon [0,\infty)\times\mathbb{R}^n\to\mathbb{R}^n$ be measurable, then we consider the equation(s)

$$dX_t = \begin{cases} dX_t^1 & = & \sum_{j=1}^n \sigma_{1j}(t,X_t)\,dB_t^j & + & b_1(t,X_t)\,dt \\ \vdots & & \vdots & & \vdots \\ dX_t^n & = & \sum_{j=1}^n \sigma_{nj}(t,X_t)\,dB_t^j & + & b_n(t,X_t)\,dt \end{cases} \tag{1.7.3}$$

where $X_0 = (x_1,\dots,x_n)\in\mathbb{R}^n$. In matrix notation (1.7.3) again reads

$$dX_t = \sigma(t,X_t)\,dB_t + b(t,X_t)\,dt \quad X_0 = x.$$

Definition 1.7.1. We say that (1.7.1) (resp. (1.7.3)) has a *pathwise solution* if there exists a continuous adapted process $(X_t)_{t\geq 0}$ satisfying (1.7.1) (resp. (1.7.3)). *Pathwise uniqueness* for (1.7.1) (resp. (1.7.3)) holds, if whenever $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ are two pathwise solutions then $\mathbb{P}(X_t = Y_t;\, \forall t \geq 0) = 1$, i.e., $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ are *indistinguishable*.

Remark 1.7.2. Note that a solution $(X_t)_{t\geq 0}$ to (1.7.1) necessarily is a semimartingale. A solution to (1.7.3) accordingly is an $\mathbb{R}^n$-valued semimartingale $(X_t) = (X_t^1, \dots, X_t^n)$.

Theorem 18. *(Existence and uniqueness for SDEs, Itō. Cf. [Øks98], p.66.) Let $(B_t)_{t\geq 0}$ be an n-dimensional Brownian motion and let $(\mathcal{F}_t)$ be the generated filtration. Let $T > 0$ and let $\sigma\colon [0,T] \times \mathbb{R}^n \to \mathbb{R}^{n\times n}$ and $b\colon [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ be measurable functions satisfying*

$$|b(t,x)| + |\sigma(t,x)| \leq C(1 + |x|), \quad x \in \mathbb{R}^n,\, t \in [0,T], \tag{1.7.4}$$

for some constant C, where $|\sigma|^2 = \sum |\sigma_{ij}|^2$, and such that

$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq D|x - y| \quad x, y \in \mathbb{R}^n,\, t \in [0,T], \tag{1.7.5}$$

for some constant D. Let Z be a square integrable r.v. independent of $\mathcal{F} = \mathcal{F}_\infty$. Then the SDE

$$dX_t = \sigma(t, X_t)\, dB_t + b(t, X_t)\, dt, \quad 0 \leq t \leq T, \quad X_0 = Z,$$

has a pathwise unique solution $(X_t)_{t\geq 0}$, adapted to the filtration $(\mathcal{F}_t^Z)$ generated by Z and $(B_s)_{0\leq s\leq t}$. Furthermore $(X_t)_{t\geq 0}$ is continuous and

$$\mathbb{E}\left[\int_0^T |X_t|^2\, dt\right] < \infty.$$

An important link between SDEs and PDEs is established through an application of Itō's formula and associates a second order partial differential operator with an SDE. Using the above notation, let σ^T denote the transpose of the matrix σ and let a be the matrix $\sigma\sigma^T$. On $\mathcal{C}^{1,2}([0,\infty) \times \mathbb{R}^n)$ we define the second order partial differential operator

$$\tilde{\mathcal{L}}_t = \frac{\partial}{\partial t} + \mathcal{L}_t = \frac{\partial}{\partial t} + \frac{1}{2}\sum_{i,j=1}^n a_{ij}(t,x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(t,x)\frac{\partial}{\partial x_i}. \tag{1.7.6}$$

Proposition 1.7.3. *(Cf. [Bas98], p.5.) Let $(X_t)_{t\geq 0}$ be a solution of (1.7.3) with σ and b bounded and measurable and let $f \in \mathcal{C}^{1,2}([0,\infty)\times\mathbb{R}^n)$. Then*

$$f(t, X_t) = f(0, X_0) + M_t + \int_0^t \widetilde{\mathcal{L}}_s f(s, X_s)\, ds, \tag{1.7.7}$$

where

$$M_t = \int_0^t \sum_{i,j=1}^n \frac{\partial}{\partial x_i} f(s, X_s)\sigma_{ij}(s, X_s)\, dB_s^j \tag{1.7.8}$$

is a martingale.

Proof. Since the components of B_t are independent, we have $d\langle B^k, B^l\rangle_t = \delta_{kl}\, dt$ and hence

$$\begin{aligned} d\langle X_i, X_j\rangle_t &= \sum_k \sum_l \sigma_{ik}(t, X_t)\sigma_{jl}(t, X_t)\, d\langle B^k, B^l\rangle_t = \\ &= \sum_k \sigma_{ik}(t, X_t)\sigma_{kj}^T(t, X_t)\, dt = a_{ij}(t, X_t)\, dt \end{aligned}$$

Itō's formula, applied to $f(t, X_t)$, yields

$$\begin{aligned} f(t, X_t) &= f(0, X_0) + \int_0^t \frac{\partial}{\partial s} f(s, X_s)\, ds + \sum_i \int_0^t \frac{\partial}{\partial x_i} f(s, X_s)\, dX_s^i + \\ &\quad + \frac{1}{2}\int_0^t \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} f(s, X_s)\, d\langle X^i, X^j\rangle_s = \\ &= f(0, X_0) + M_t + \int_0^t \frac{\partial}{\partial s} f(s, X_s)\, ds + \sum_i \int_0^t \frac{\partial}{\partial x_i} f(s, X_s) b_i(s, X_s)\, ds + \\ &\quad + \frac{1}{2}\int_0^t \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} f(s, X_s) a_{ij}(s, X_s)\, ds = \\ &= f(0, X_0) + M_t + \int_0^t \widetilde{\mathcal{L}}_s f(s, X_s)\, ds. \end{aligned}$$

□

Remark 1.7.4. We will call a process $(X_t)_{t\geq 0}$ and an operator $\widetilde{\mathcal{L}}_t$ *associated* if $(X_t)_{t\geq 0}$ satisfies (1.7.3) for $\widetilde{\mathcal{L}}_t$ given by (1.7.6) and $a = \sigma\sigma^T$. We call σ (resp. a) the *diffusion coefficient* of $(X_t)_{t\geq 0}$ (resp. of $\widetilde{\mathcal{L}}_t$) and b the drift coefficient.

Remark 1.7.5. Another notion to associate PDE theory to SDEs is the *infinitesimal generator* of a diffusion which does not associate a PDE to an SDE, but a linear second order partial differential operator, which is exactly the operator $\widetilde{\mathcal{L}}_t$ above.

(Cf. [Øks98], p.117ff.) We briefly state the definitions and some results in the time-homogeneous case.

Definition 1.7.6. (Cf. [Øks98], p.117.) Let $(X_t)_{t\geq 0}$ be the solution of an SDE of the (time-homogeneous) form of (1.7.3) satisfying the assumptions of Theorem 18. Then we call $(X_t)_{t\geq 0}$ an Itō-diffusion and the (infinitesimal) generator A of $(X_t)_{t\geq 0}$ is given by

$$Af(x) = \lim_{t\searrow 0} \frac{\mathbb{E}^x[f(X_t)] - f(x)}{t} \quad x \in \mathbb{R}^n. \tag{1.7.9}$$

We denote the set of functions $f\colon \mathbb{R}^n \to \mathbb{R}$ such that the above limit exists for x by $\mathcal{D}_A(x)$ while the set of functions where the limit exists for all $x \in \mathbb{R}^n$ is denoted by $\mathcal{D}_A$.

Proposition 1.7.7. *(Cf. [Øks98], p.119.) Let $(X_t)_{t\geq 0}$ be a (time-homogeneous) Itō-diffusion given by the SDE (1.7.3). Let $f \in \mathcal{C}_0^2(\mathbb{R}^n)$, then $f \in \mathcal{D}_A$ and the infinitesimal generator A is given by*

$$Af(x) = \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} f(x) + \sum_{i=1}^{n} b_i(x)\frac{\partial}{\partial x_i} f(x). \tag{1.7.10}$$

Proof. Note that we are in the (time-homogeneous) setting of Proposition 1.7.3, hence we can compute the expectation of (1.7.7) w.r.t. $\mathbb{P}^x$ and use Fubini.

$$\begin{aligned}
\mathbb{E}^x[f(X_t)] &= \mathbb{E}[f(X_0)] + \mathbb{E}[M_t] + \mathbb{E}[\int_0^t \widetilde{\mathcal{L}} f(X_s)\, ds] = \\
&= f(x) + 0 + \mathbb{E}[\int_0^t \widetilde{\mathcal{L}} f(X_s)\, ds] = \\
&= f(x) + \int_0^t \mathbb{E}[\widetilde{\mathcal{L}} f(X_s)]\, ds.
\end{aligned}$$

Hence

$$\begin{aligned}
Af(x) &= \lim_{t\searrow 0} \frac{\mathbb{E}^x[f(X_t)] - f(x)}{t} = \\
&= \frac{\partial}{\partial t}\left(\int_0^t \mathbb{E}^x[\widetilde{\mathcal{L}} f(X_s)]\, ds\right)\Big|_{t=0} = \\
&= \widetilde{\mathcal{L}} f(x).
\end{aligned}$$

□

Remark 1.7.8. To extend the definition of the infinitesimal generator to the time dependent case, we consider a whole family of generators

$$L_t f(x) = \lim_{s \searrow 0} \frac{\mathbb{E}[f(X_{t+s}) \,|X_t = x] - f(x)}{s}, \quad t \geq 0,\, x \in \mathbb{R}^n. \tag{1.7.11}$$

Then, we define

$$\widetilde{L}_t f(t,x) = \frac{\partial}{\partial t} f(t,x) + \Big(L_t f(t,\cdot)\Big)(x),$$

and $\mathcal{D}_{\widetilde{L}_t}$ as the set of functions, for which the above limit exists for all $t \geq 0$ and $x \in \mathbb{R}^n$. Note, that this is exactly the operator we were dealing with in Proposition 1.7.3.

From the above we see that, if $(X_t)_{t\geq 0}$ is an Itō-diffusion in $\mathbb{R}^n$ with generator A, $f \in C_0^2(\mathbb{R}^n)$ and if we set

$$u(t,x) = \mathbb{E}^x[f(X_t)],$$

then $u(t,x)$ is differentiable in t and its derivative is given by

$$\frac{\partial u}{\partial t} = \mathbb{E}^x[Af(X_t)].$$

One of the most powerful links between the analytic theory of PDEs and stochastic calculus is given by the *Kolmogorov backward equation*, which turns, if a killing term is involved, into the *Feynman-Kac formula*.

Theorem 19. *(Kolmogorov's backward equation, cf. [Øks98], p.133) Let $(X_t)_{t\geq 0}$ be an Itō-diffusion and let $f \in \mathcal{C}_0^2(\mathbb{R}^n)$. Define*

$$u(t,x) = \mathbb{E}[f(X_t)],$$

then
(i) $u(t,\cdot) \in \mathcal{D}_A$ for each t and

$$\frac{\partial u}{\partial t} = Au, \quad t > 0,\, x \in \mathbb{R}^n, \tag{1.7.12}$$

$$u(0,x) = f(x) \quad x \in \mathbb{R}^n, \tag{1.7.13}$$

where Au is understood as A applied to the mapping $x \mapsto u(t,x)$.

(ii) If $w(t,x) \in \mathcal{C}^{1,2}(\mathbb{R} \times \mathbb{R}^n)$ is a bounded function satisfying the conditions of (i), then $w(t,x) = u(t,x)$.

Remark 1.7.9. Often the backward equation, especially the Feynman-Kac formula, is not stated for $t > 0$ and an initial value $u(0, x) = f(x)$, but for a finite time horizon $t \in [0, T]$ and as a terminal value problem with $u(T, x) = g(x)$. Because of the time inversion in this case the equation reads

$$-\frac{\partial u}{\partial t} = Au, \quad 0 \leq t < T,\, x \in \mathbb{R}^n, \tag{1.7.14}$$

$$u(T, x) = g(x) \quad x \in \mathbb{R}^n. \tag{1.7.15}$$

We then obtain for $g \in \mathcal{C}_0^2(\mathbb{R}^n)$ the unique solution to the above equation as

$$u(t, x) = \mathbb{E}^x[g(X_{T-t})]. \tag{1.7.16}$$

(For the case of the multidimensional Feynman-Kac formula for Brownian motion cf. for instance [KS88], p.268.)

Before we turn to the type of solutions given via the martingale problem, we state a theorem of Engelbert and Schmidt which gives a criterion for the existence of solutions for one-dimensional time-homogeneous SDEs without drift (even for coefficients which are not Lipschitz).

Theorem 20. *(Engelbert-Schmidt, cf. [KS88], p.335.) Let $(B_t)_{t\geq 0}$ be a linear Brownian motion and let $\sigma\colon \mathbb{R} \to \mathbb{R}$ be a Borel-measurable function. Define*

$$dX_t = \sigma(X_t)\, dB_t, \quad X_0 \sim \mu. \tag{1.7.17}$$

Then, for every initial distribution μ, this equation has a solution which is unique in the sense of finite dimensional distributions if and only if

$$\{x \in \mathbb{R} \,|\, \int_{-\varepsilon}^{\varepsilon} \frac{dy}{\sigma(x+y)^2} = \infty \quad \forall \varepsilon > 0\} = \{x \in \mathbb{R} \,|\, \sigma(x) = 0\}. \tag{1.7.18}$$

Example 1.7.10. Consider the (simple) case, where $\sigma(x) = 1/x$ (which we will need in Section 2.4). Then $\{x \in \mathbb{R} \,|\, \sigma(x) = 0\} = \emptyset$. On the other hand

$$\int_{-\varepsilon}^{\varepsilon} \frac{dy}{\sigma(x+y)^2} = \int_{-\varepsilon}^{\varepsilon} (x+y)^2\, dy = \frac{(x+\varepsilon)^3 - (x-\varepsilon)^3}{3} < \infty \quad \forall x \in \mathbb{R},\, \varepsilon > 0.$$

This shows that also the left hand side of (1.7.18) is equal to $\emptyset$, hence the condition is satisfied.

1.7.2 The martingale problem

This approach to solve stochastic differential equations, due to Stroock and Varadhan, is essentially based upon the observation we made in Proposition 1.7.3. There we were given a solution $(X_t)_{t\geq 0}$ of an SDE and identified the associated operator $\widetilde{\mathcal{L}}_t$. Then we observed that, for $f \in \mathcal{C}^{1,2}([0,\infty) \times \mathbb{R}^n)$, the process

$$M_t = f(t, X_t) - f(0, X_0) - \int_0^t \widetilde{\mathcal{L}}_s f(s, X_s)\, ds, \tag{1.7.19}$$

is a martingale. Stroock and Varadhan take the opposite point of view. One is given an operator $\mathcal{L}$ on a space of (test) functions and asks for a process $(X_t)_{t\geq 0}$ such that

$$M_t := f(X_t) - \int_0^t \mathcal{L} f(X_s)\, ds$$

is a martingale for each f in the domain $D(\mathcal{L})$ of $\mathcal{L}$. I.e., one looks for solutions to the SDE via the associated differential operator. The main advantage of this approach, compared to the Itō-theory where the drift and diffusion coefficient have to be Lipschitz, consists in weaker assumptions necessary for existence and uniqueness of a solution to the martingale problem for $\mathcal{L}$. However, uniqueness in this case refers to uniqueness in the sense of finite dimensional distributions, not pathwise uniqueness.

First, we note that there are several (equivalent) ways to state the martingale problem and its solution. We start with a general definition of the martingale problem with polish state space $\mathcal{E}$ and operators $\mathcal{L}$ acting on the set of bounded continuous functions $\mathcal{C}_b(\mathcal{E})$ with domain $D(\mathcal{L})$. Nevertheless it turns out, that the interesting case is the one of (elliptic/parabolic) second order partial differential operators acting on smooth or $\mathcal{C}_b^{1,2}([0,\infty) \times \mathbb{R}^n)$ functions which is, of course, covered by this general setting.

Definition 1.7.11. Let $\mathcal{L}$ be an operator on $\mathcal{C}_b(\mathcal{E})$ with domain $D(\mathcal{L})$. A measurable process $(X_t)_{t\geq 0}$ adapted to $(\mathcal{F}_t)$ is said to be a solution to the martingale problem for $(\mathcal{L}, \mu)$ w.r.t. $(\mathcal{F}_t)$ if, for all $f \in D(\mathcal{L})$,

$$M(t) := f(X(t)) - \int_0^t \mathcal{L} f(X(s)) ds$$

is an $(\mathcal{F}_t)$-martingale and $\mathbb{P} \circ X_0^{-1} = \mu$. (If $(\mathcal{F}_t)$ is the natural filtration, it can be dropped from the statement.)

Since uniqueness for solutions of the martingale problem is only meant w.r.t. the f.d.d.s, it is natural to consider the process $(X_t)_{t\geq 0}$, started at x, as probability measure $\widetilde{\mathbb{P}}^x$ on the path space $((\mathbb{R}^n)^{[0,\infty)}, \mathcal{B}(\mathbb{R}^n)^{[0,\infty)})$. Since we restrict ourselves to continuous functions we consider measures on $(\mathcal{C}([0,\infty),\mathbb{R}^n),\mathcal{B})$.

Definition 1.7.12. Let $\mathcal{L}_t := \frac{1}{2}\sum_{i,j=1}^n a_{ij}(t,x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(t,x)\frac{\partial}{\partial x_i}$ be a second order linear partial differential operator. Let $\omega(t)$ be the coordinate mappings on $\mathcal{C}([0,\infty),\mathbb{R}^n)$. A measure $\mathbb{P}^x$ is said to be a solution to the martingale problem for $(\mathcal{L}_t,\delta_x)$, if under $\mathbb{P}^x$ the coordinate process $(\omega(t))_t$ is a solution to the martingale problem.

The martingale problem is said to be *well posed*, if there is a unique measure $\mathbb{P}^x$ which solves the martingale problem on $(\mathcal{C}([0,\infty),\mathbb{R}^n),\mathcal{B},(\mathcal{B}_t))$, where $\mathcal{B}_t$ is defined as the Borel σ-algebra on $(\mathcal{C}([0,t],\mathbb{R}^n)$.

If the martingale problem is well posed for every x in $\mathbb{R}^n$ and μ is a probability measure on the state space $\mathbb{R}^n$, then the martingale problem for $(\mathcal{L}_t,\mu)$ is said to be well posed, if there is a unique measure $\mathbb{P}^\mu$ on $(\mathcal{C}([0,\infty),\mathbb{R}^n),\mathcal{B},(\mathcal{B}_t))$ that solves the martingale problem for $(\mathcal{L}_t,\mu)$.

The important link between an SDE and the martingale problem for its infinitesimal generator $\mathcal{L}$ is given by the following theorem.

Theorem 21. *(Cf. [Bas98], p.98.) Let*

$$dX(t) = \sigma(X_t)dW(t) + b(X_t)dt, \qquad X_0 = x_0, \tag{1.7.20}$$

be a stochastic differential equation. Set $a = \sigma\sigma^T$, and let

$$\mathcal{L} := \frac{1}{2}\sum_{i,j=1}^n a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(x)\frac{\partial}{\partial x_i}$$

be the associated uniformly elliptic operator with bounded measurable coefficients a_{ij} and b_i. Then uniqueness in the f.d.d.s for (1.7.20) holds if and only if the martingale problem for $(\mathcal{L},\delta_{x_0})$ is well posed.
Weak existence for (1.7.20) holds if and only if there exists a solution to the martingale problem for $(\mathcal{L},\delta_{x_0})$.

Remark 1.7.13. The above theorem is also true in the time-inhomogeneous case. See [SV79, Chapter 8.] for statement and proof of the above equivalence in the time dependent setting.

Well posedness in the time-dependent setting is defined as follows.

Definition 1.7.14. For a probability measure μ on the state space E, the martingale problem for $((\mathcal{L}_t), \mu)$ is well posed, if whenever X and Y are solutions to the martingale problem for $(\mathcal{L}_t)$, defined respectively on $(\Omega, \mathcal{F}, \mathbb{P})$ and $(\Gamma, \mathcal{G}, \mathbb{Q})$ w.r.t. some filtration and satisfying $\mathbb{P} \circ X_0^{-1} = \mathbb{Q} \circ Y_0^{-1} = \mu$ we have

$$\mathbb{P}(X_t \in U) = \mathbb{Q}(Y_t \in U) \quad \forall t > 0 \text{ and } \forall U \text{ Borel in } E.$$

The question, however, is under which conditions on the coefficients of the operator $\mathcal{L}$ we get a unique solution for the martingale problem. A celebrated result of Stroock/Varadhan gives sufficient conditions for well-posedness.

Theorem 22. *Let*

$$\mathcal{L}_t = \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(t,x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(t,x) \frac{\partial}{\partial x_i},$$

where $b(t,x) : [0,\infty) \times \mathbb{R}^n \to \mathbb{R}^n$ and $a(t,x) : [0,\infty) \times \mathbb{R}^n \to S_d^+$ are bounded and measurable, where S_d^+ is the set of nondegenerate elements of the space of symmetric nonnegative definite $n \times n$ real matrices. Assume that, for each $T > 0$ and $x \in \mathbb{R}^n$,

$$\inf_{0 \le s \le T} \inf_{z \in \mathbb{R}^n \setminus \{0\}} \langle z, a(s,x) z \rangle > |z|^2 \quad \text{and}$$
$$\lim_{y \to x} \sup_{0 \le s \le T} \|a(s,y) - a(s,x)\| = 0, \qquad x \in \mathbb{R}^n.$$

Then the martingale problem for a and b is well posed.

Proof. See [SV79, Thm. 7.2.1] □

Remark 1.7.15. Since the above theorem is stated just in terms of the coefficients a and b, the question arises, whether there is a difference if the martingale problem is stated for $\mathcal{L}_t$ or for $(\frac{\partial}{\partial t} + \mathcal{L}_t)$. The equivalence of the two problems is also proved by Stroock and Varadhan who show that, for time dependent and even nondeterministic coefficients, several formulations of a martingale are equivalent. We are interested only in two of six possible formulations, stated below.

Theorem 23. *Let $b(t,\omega) : [0,\infty) \times \Omega \to \mathbb{R}^n$ and $a(t,\omega) : [0,\infty) \times \Omega \to S_d$ be bounded progressively measurable functions, let $\xi(t,\omega) : [0,\infty) \times \Omega \to \mathbb{R}^n$ be progressively* *measurable and right continuous in t; for $f \in \mathcal{C}^2(\mathbb{R}^n)$ define*

$$(\mathcal{L}_t(\omega)f)(x) := \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(t,\omega)\frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(t,\omega)\frac{\partial f}{\partial x_i}.$$

Then TFAE:

$$\begin{aligned}
&(i)\ f(\xi_t) - \int_0^t (\mathcal{L}_u f)(\xi_u)\,du \ \textit{ is a martingale relative to}\\
&\quad (\Omega, \mathcal{F}_t, \mathbb{P}) \ \textit{ for } t \geq 0, \ \textit{ for all } f \in \mathcal{C}_0^\infty(\mathbb{R}^n).\\
&(ii)\ f(t,\xi_t) - \int_0^t (\frac{\partial}{\partial u} + \mathcal{L}_u) f(u,\xi_u)\,du \ \textit{ is a martingale relative to}\\
&\quad (\Omega, \mathcal{F}_t, \mathbb{P}) \ \textit{ for } t \geq 0, \ \textit{ for all } f \in \mathcal{C}_b^{1,2}([0,\infty)\times\mathbb{R}^n).
\end{aligned}$$

Remark 1.7.16. Clearly this works also in the case of deterministic coefficients. (Cf. the proof of [SV79, Thm. 4.2.1.]). Hence $(\xi_t)_t$ is a solution to the martingale problem for $\mathcal{L}_t$ if and only if it is a solution for $(\frac{\partial}{\partial t} + \mathcal{L}_t)$, i.e., the martingale problem is solely determined by the drift and diffusion coefficients.

Remark 1.7.17. What is of particular interest in our current inquiry is the fact that solutions to the martingale problem are (up to versions) uniquely determined via their one dimensional marginals. For a proof in the time-homogeneous case cf. [KS88, Proposition 5.4.27], p.326 or [EK86, Theorem 4.2], p.184.

1.8 First steps towards mimicking

Definition 1.8.1. Two processes X and Y are said to be equal in the k-dimensional marginal distributions if, for any $t_1, \ldots, t_k$; $t_i \in \mathbb{R}^+$ and $A_i \in \mathcal{B}(\mathbb{R}^d)$, $1 \leq i \leq k$,

$$\mathbb{P}(X_{t_1} \in A_1, \ldots, X_{t_k} \in A_k) = \mathbb{P}(Y_{t_1} \in A_1, \ldots, Y_{t_k} \in A_k). \tag{1.8.1}$$

This will be denoted by $X \overset{(k.d)}{=} Y$.

If two random variables X and Y are identically distributed we frequently will use the notation

$$X \overset{(law)}{=} Y. \tag{1.8.2}$$

Definition 1.8.2. Let X be a stochastic process. If another process $\widetilde{X}$ is equal to X in all one-dimensional marginals, we call $\widetilde{X}$ a *mimicking process.*

Remark 1.8.3. In the case of Brownian motion we will call a mimicking process X *fake Brownian motion* if $X_t \overset{(law)}{=} B_t$ for all $t \geq 0$, but X is *not* Brownian motion itself. (Cf. [Ole08].)

A key concept in studying the relationship between martingales and one dimensional marginals is the so called convex order. We already observed that for a martingale X and any convex function ψ the process $(\psi(X_t))_{t\geq 0}$ is a submartingale. However, convex functions also play a crucial role in the setup of martingales themselves, precisely in the ordering of the one-dimensional marginals.

Definition 1.8.4. Let X and Y be two real valued random variables. X is said to dominate Y in the convex order if for every convex function $\psi\colon \mathbb{R} \to \mathbb{R}$ (provided $\mathbb{E}[|\psi(X)|] < \infty$ and $\mathbb{E}[|\psi(Y)|] < \infty$) we have

$$\mathbb{E}[\psi(X)] \geq \mathbb{E}[\psi(Y)].$$

We denote this order relation by

$$X \overset{(c)}{\geq} Y.$$

Remark 1.8.5. It is clear from the definition, that this order relation does not depend on the values of the random variables, but just on the distributions. Properly speaking, it is an order relation between probability measures (on the real line).

Definition 1.8.6. A process $(X_t)_{t\geq 0}$ is said to be increasing in the convex order, if for every $s \leq t$, $X_s \overset{(c)}{\leq} X_t$.

Proposition 1.8.7. *A martingale $(M_t)_{t\geq 0}$ is increasing in the convex order.*

Proof. Let $\psi\colon \mathbb{R} \to \mathbb{R}$ be a convex function, and let $s < t$, then by Jensen's inequality and iterated conditioning we obtain that

$$\begin{aligned}\mathbb{E}[\psi(X_t)] = \mathbb{E}[\mathbb{E}[\psi(X_t) \mid \mathcal{F}_s]] \geq \mathbb{E}[\psi(\underbrace{\mathbb{E}[X_t \mid \mathcal{F}_s]}_{X_s})] = \\ = \mathbb{E}[\psi(X_s)].\end{aligned}$$

□

So, a family of random variables necessarily has to increase in the convex order in order to form a martingale. The question whether this condition is sufficient was answered affirmatively by V. Strassen ([Str65]) for a countable family $(\mu_n)_{n\in\mathbb{N}}$ of probability measures on $\mathbb{R}$. Strassen proved that, given such a sequence increasing in the convex order, there exists a markov martingale $(M_n)_{n\geq 0}$ with these marginals. J. Doob established the continuous time case but only for compact state space and omitting the Markov property. The definitive answer to the question of sufficiency was given by H.G. Kellerer, who proved (a more general version of) the following

Theorem 24. *(H.G. Kellerer, cf [Kel72])*
Let $(\mu_t)_{t\in\mathbb{R}^+}$ be a family of probability measures on $\mathbb{R}$. Let $(X_t)_{t\geq 0}$ be a family of real valued random variables. $X_t \sim \mu_t$, $\mathbb{E}[X_t] < \infty$ for all t and $X_s \overset{(c)}{\leq} X_t$ for all $s < t$. Then there exists a measure $\mathbb{P}_X$ on the canonical space $(\mathbb{R}^{[0,\infty)}, \mathcal{B}(\mathbb{R})^{[0,\infty)})$ with these one dimensional marginals s.t. $(X_t)_{t\geq 0}$ is a martingale.

Proof. Cf. [Kel72]. □

Remark 1.8.8. There is a slight difference in the notion of convex order used here and used by Kellerer. Kellerer defines the order relation on the probability measures not via all convex functions, but via monotone increasing convex functions. In our formulation, the convex order already implies that all μ_t have the same constant mean which Kellerer has to assume additionally to obtain a martingale, since in general he obtains just a submartingale. But on the other hand the process in his construction can be chosen to be Markov.

2 Mimicking Brownian Motion

In this chapter we study in detail how Brownian Motion can be mimicked by processes having the same one dimensional marginals. Using a notion of Oleszkiewicz we call such processes *fake Brownian motions*.

Section 2.1 recalls equivalent definitions and some properties of Brownian motion.

In Section 2.2 we take a look at continuous semimartingales that not only have the same marginals but also the same quadratic variation as Brownian motion; however, they are not martingales. (The material of Section 2.2 is taken from a paper of Föllmer, Wu and Yor, [FWY00].)

Section 2.3 discusses a 2006/07 result by Hamza/Klebaner and gives a construction of a discontinuous fake Brownian motion which is a martingale and enjoys the Markov property. In 2007, it was still an open question if there existed a *continuous* fake Brownian motion which is a martingale.

In Section 2.4 we present the results of Albin, [Alb08], and Oleszkiewicz, [Ole08] who positively answered the above question and gave explicit constructions of continuous martingale fake BMs which are, nota bene, not Markov.

Finally, in Section 2.5 we add the Markov property; we first will see that a continuous Markov martingale which has Brownian marginals and is adapted to the Brownian filtration is already Brownian motion itself. Secondly, as a special case of a result in [Low08a], we will see that any continuous Strong Markov martingale with Brownian marginals is already Brownian motion.

2.1 Essential Facts concerning Brownian Motion

We briefly recall some (characteristic) properties of Brownian motion. First of all, there are at least four equivalent definitions of linear Brownian motion. The first one was already given in Section 1.2.

Definition 2.1.1. Let $(B_t)_{t\geq 0}$ be a real valued stochastic process on $(\Omega, \mathcal{F}, \mathbb{P})$. $(B_t)_{t\geq 0}$ is called a *(standard) Brownian Motion* iff

(i) $B_t \sim \mathcal{N}(0,t)$

(ii) $\forall t_0 \leq \ldots \leq t_k; t_i \in \mathbb{R}^+$ and for $0 \leq i \leq k-1$: $(B_{t_{i+1}} - B_{t_i})$ are independent random variables.

(iii) $(B_t)_{t\geq 0}$ is a.s. continuous.

Equivalently Brownian motion can be defined as a Gaussian process with a certain covariance structure.

Definition 2.1.2. Let $(B_t)_{t\geq 0}$ be a real valued stochastic process on $(\Omega, \mathcal{F}, \mathbb{P})$.

(i)' (B_t) is a centered Gaussian process.

(ii)' $\mathrm{Cov}(B_s, B_t) = s \wedge t$

(iii)' $(B_t)_{t\geq 0}$ is a.s. continuous.

Proposition 2.1.3. *The above definitions are equivalent.*

Proof. Clearly (i) and (ii) imply that $(B_t)_{t\geq 0}$ is a centered Gaussian process. To see that (i) and (ii) imply (ii)' let $s < t$, then $(B_t - B_s)$ is independent of B_s and

$$\mathbb{E}[B_s B_t] = \mathbb{E}[B_s(B_t - B_s) + B_s^2] = \mathbb{E}[B_s]\mathbb{E}[B_t - B_s] + \mathbb{E}[B_s^2] = 0 + s = s.$$

Because of the symmetry in t and s, $\mathrm{Cov}(B_s, B_t) = s \wedge t$.

To prove the other direction let $0 = t_0 < t_1 < \cdots < t_k$ and consider the random vector $X = (B_{t_1}, \ldots, B_{t_k})$ and the vector of the increments $Y = (B_{t_1} - B_{t_0}, \ldots, B_{t_k} - B_{t_{k-1}})$. Then, if M is the $k \times k$ matrix with diagonal entries 1, -1 below and 0 elsewhere, $Y = MX$. If $D := (t_i \wedge t_j)_{1\leq i,j\leq k}$ then $MDM^T = C$ where C is a diagonal matrix with entries $t_j - t_{j-1}$ for $1 \leq j \leq k$. Now X has a centered multivariate normal distribution with covariance matrix D if and only if $Y = MX$ has a centered multivariate normal distribution with covariance matrix C. Hence we see, that the increments $B_{t_1} - B_{t_0}, \ldots, B_{t_k} - B_{t_{k-1}}$ are independent and have law $\mathcal{N}(0, t_j - t_{j-1})$, in particular $B_t \sim \mathcal{N}(0,t)$. □

The third (possible) definition is usually stated as a characterization-theorem, known as Lévy's martingale characterization. We state it for d-dimensional Brownian motion

Theorem 25. *(P. Lévy 1948; cf. [KS88], p.157)*
Let $X = (X_t^1, \dots, X_t^d)$ be a continuous, $(\mathcal{F}_t)_{t\geq 0}$-adapted process in $\mathbb{R}^d$ such that for every component $1 \leq k \leq d$ the process

$$M_t^k := X_t^k - X_0^k, \quad t \in [0, \infty)$$

is a continuous local martingale relative to $(\mathcal{F}_t)_{t\geq 0}$, and the cross-variations are given by

$$\langle M^i, M^j \rangle_t = \delta_{ij} t, \quad 1 \leq i, j \leq d, \tag{2.1.1}$$

then $((X_t)_{t\geq 0}, (\mathcal{F}_t)_{t\geq 0})$ is a d-dimensional Brownian motion.

Proof. cf. [KS88], p.157. □

Remark 2.1.4. The assumption of continuity is essential within this characterization, since the compensated Poisson process with intensity $\lambda = 1$ provides an example of a square integrable martingale with quadratic variation $\langle M \rangle_t = t$. But this process is *discontinuous*. Moreover, what is remarkable about this theorem is that no assumptions on the distributions are made, similar to Theorem 36 below which asserts that any continuous martingale is a time-changed Brownian motion. In some sense the path-properties encode the information w.r.t. the distributions, or put otherwise, the continuity of the paths ensures uniqueness of the f.d.d.s. In mimicking Brownian motion and martingales, we also will encounter the crucial role of continuous paths in questions of uniqueness.

The fourth definition characterizes Brownian motion as a certain Lévy process.

Theorem 26. *Let $(X_t)_{t\geq 0}$ be a stochastic process on $(\Omega, \mathcal{F}, \mathbb{P})$ such that*

(i) *$(X_t)_{t\geq 0}$ has stationary independent increments.*
(ii) *$(X_t)_{t\geq 0}$ is a.s. continuous and $X_0 = 0$.*
(iii) *All increments $X_t - X_s$, where $s \leq t$, have mean zero.*

Then $(X_t)_{t\geq 0}$ is a standard Brownian motion.

Proof. Cf. [Kni81], p.15. □

In order to distinguish Brownian motion from its mimicking processes it is useful to know some further properties of BM.

Proposition 2.1.5. *(Cf. [RY99], p.19.)*
Let $(B_t)_{t\geq 0}$ be a standard Brownian motion, then the following properties hold.

(i) (time-homogeneity) For any $s > 0$, the process $(B_{t+s} - B_s)_{t\geq 0}$ is a standard BM independent of $\sigma(B_u, u \leq s)$.
(ii) (symmetry) The process $(-B_t)_{t\geq 0}$ is a standard BM.
(iii) (scaling) For every $c > 0$, the process $(cB_{t/c^2})_{t\geq 0}$ is a standard BM.
(iv) (time inversion) The process X defined by $X_0 = 0$, $X_t = tB_{1/t}$ for $t > 0$ is a standard BM.

Proof. Clearly in (i), (ii) and (iii) the paths remain continuous, since only continuous transformations are applied. To see that in all four cases the increments are properly distributed is an easy exercise, e.g. in the case (iii). Let $0 = t_0 < t_1 < \cdots < t_k$, then each increment $cB_{t_j/c^2} - cB_{t_{j-1}/c^2} \sim c\mathcal{N}(0, (t_j - t_{j-1})/c^2) = \mathcal{N}(0, t_j - t_{j-1})$ for $1 \leq j \leq k$ and clearly they are independent.
What remains to show is that in (iv) the process X is continuous in 0, since it is obviously continuous on $(0, \infty)$. However, for $t \in (0, \infty)$, (X_t) and (B_t) are versions of each other and since $\lim_{t\to 0, t\in\mathbb{Q}} B_t = 0$ it follows that $\lim_{t\to 0, t\in\mathbb{Q}} X_t = 0$ a.s. Because X is continuous on $(0, \infty)$, we get $\lim_{t\to 0, t\in\mathbb{R}^+} X_t = 0$. □

Remark 2.1.6. Property (iv) implies the Strong Law of Large Numbers for Brownian motion.
$$\mathbb{P}\left(\lim_{t\to\infty} \frac{B_t}{t} = 0\right) = 1.$$

Another power- and beautiful result is the Law of the Iterated Logarithm.

Theorem 27. *(Cf. [RY99], p.56ff.) Let $(B_t)_{t\geq 0}$ be a standard Brownian motion, then*

$$\mathbb{P}\left(\overline{\lim}_{t\to\infty} \frac{B_t}{\sqrt{2t\log(\log t)}} = 1\right) = \mathbb{P}\left(\underline{\lim}_{t\to\infty} \frac{B_t}{\sqrt{2t\log(\log t)}} = -1\right) = 1. \tag{2.1.2}$$

As we already saw in Section 1.4 (by writing down the transition function) Brownian motion is a *Markov process* and it enjoys also the strong Markov property. This can be seen by checking that the transition function is a Feller semi-group.

Of course there are myriads of other interesting properties of Brownian motion such as the non-differentiability of the paths, but few of them serve to characterize it. If necessary, we will notice some of them at the appropriate place. Now we start the proper business of mimicking.

2.2 Quadratic Variation & Marginals – continuous

First of all, recall the two examples of mimicking processes we encountered in the introduction. In the case of $X_t := \sqrt{t}\, N$ no assumptions were made except matching marginals. Nonetheless it turned out that (X_t) has continuous, even differentiable paths on $(0, \infty)$ hence zero quadratic variation. Precisely because of the zero quadratic variation of (X_t) in the second example we considered a continuous process (Y_t) on the unit interval with nonzero quadratic variation to get somewhat closer to Brownian motion. (Y_t) was constructed via a proper Brownian motion and a scaled Brownian motion, having quadratic variation $\langle Y \rangle_t = t$ for $t \leq \frac{1}{2}$ and $\langle Y \rangle_t = (\sqrt{2}-1)^2\, t$ for $t > \frac{1}{2}$. But (Y_t) is neither a martingale nor a semimartingale.

Next we could ask if there exists a fake Brownian motion (X_t) having quadratic variation $\langle X \rangle_t = t$? However, if we additionally ask for continuous paths we know from Lévy's characterization that the resulting fake Brownian motion cannot be a martingale, otherwise it would be Brownian motion itself. So, to get a (proper) fake Brownian motion with quadratic variation t we either have to drop the path property and consider discontinuous processes or we drop the martingale property. In this section we will discuss the latter option and present parts of the paper *On weak Brownian motions of arbitrary order* by Föllmer, Wu and Yor ([FWY00]). In the next section, however, we shall consider discontinuous fake Brownian motions which are martingales but have quadratic variation not equal to t. Whether there exists a discontinuous fake Brownian motion which is a martingale and has quadratic variation t is unknown to the author (and, as he supposes, to many others as well).

The main reason why it is convenient to consider fake Brownian motions (X_t) with paths having quadratic variation $\langle X \rangle_t = t$ lies in the fact that it allows to apply Itō-calculus in a strictly pathwise manner as proved by Föllmer in [Föl81]. In this case the Itō-integral

$$\int_0^t f(X_s)\, dX_s$$

exists as a pathwise limit of non-anticipating Riemann sums along dyadic partitions for any bounded $f \in \mathcal{C}^1(\mathbb{R})$ (cf. [FWY00], p.449). In particular it satisfies Itō's formula and, for any bounded $f \in \mathcal{C}^1(\mathbb{R})$,

$$\mathbb{E}\left[\int_0^t f(X_s)\,dX_s\right] = 0. \tag{2.2.1}$$

To see this consider a continuous fake Brownian motion on the unit interval with coordinate process $(X_t)_{t\in[0,1]}$ (cf. [FWY00], p.451). Denote by $\widetilde{\mathbb{P}}$ its law on $\mathcal{C}[0,1]$ and assume that $\widetilde{\mathbb{P}}$ is concentrated on the set of continuous paths which have quadratic variation $\langle X\rangle_t = t$ along the sequence of dyadic partitions. Let $f \in C^1(\mathbb{R})$ be bounded and F its antiderivative, i.e. $F' = f$, then

$$\int_0^t f(X_s)\,dX_s = F(X_t) - F(X_0) - \frac{1}{2}\int_0^t f'(X_s)\,ds.$$

By Fubini's Theorem we get that

$$\widetilde{\mathbb{E}}\left[\int_0^t f(X_s)\,dX_s\right] = \widetilde{\mathbb{E}}[F(X_t)] - \widetilde{\mathbb{E}}[F(X_0)] - \frac{1}{2}\int_0^t \widetilde{\mathbb{E}}[f'(X_s)]\,ds$$

depends only on the one dimensional marginals of X. Since the Brownian marginals are centered, we get 2.2.1 just as in the case of the Itō-integral.

So far we still have to *assume* the existence of a fake Brownian motion with quadratic variation t. In Theorem 29 below we will prove the existence of continuous fake Brownian motions having this property. To be precise, we show the existence of continuous Gaussian semimartingales with Brownian marginals and quadratic variation t which are different from Brownian motion. (Until the end of this section we will follow very closely [FWY00, Sec.6. & Sec.7].)

2.2.1 Brownian motions and Volterra kernels

We again consider (fake) Brownian motions on the unit interval and define semimartingales with the help of integral kernels.

Definition 2.2.1. Let $l\colon [0,1]\times[0,1] \to \mathbb{R}$ be a function s.t.

$$l(u,v) = 0, \quad \text{for } 0 \le u < v \le 1,$$

and

$$\tilde{l}(u,v) := \begin{cases} l(u,v) & \text{for } u \leq v \\ l(v,u) & \text{for } u > v \end{cases}$$

is continuous on $(0,1)\times(0,1)$, then we call l a continuous Volterra kernel.

We will consider Gaussian semimartingales of the form

$$X_t := B_t - \int_0^t \int_0^u l(u,v)\, dB_v\, du. \tag{2.2.2}$$

Remark 2.2.2. Under the condition that

$$\int_0^t \left(\int_0^u l(u,v)^2\, dv \right)^{\frac{1}{2}} du < \infty, \tag{2.2.3}$$

the representation (2.2.2) is well defined as the semimartingale decomposition of X w.r.t. the filtration $(\mathcal{F}_t^B)$. From Proposition 1.6.8 we get that

$$\langle X, X\rangle_t = \langle B, B\rangle_t = t.$$

Remark 2.2.3. If we do *not* assume $l \in L^2([0,1]\times[0,1])$ the above representation is not necessarily unique. For instance

$$X_t := B_t - \int_0^t \frac{B_u}{u}\, du$$

is a proper Brownian motion and thus X has two Volterra representations. A trivial one with $l_X(u,v) \equiv 0$ and the above one, where $l_B(u,v) = 1/u$, for $v \leq u$.

If X is a proper Brownian motion admitting a Volterra representation of the form (2.2.2) then, as shown by Hitsuda (cf. [FWY00], p.474-475), the kernel l cannot be square integrable unless $l \equiv 0$. Nevertheless it is possible to characterize Brownian motions with Volterra representation even if the kernels are not square integrable. This is done in Theorem 28 below. More important for us, however, is that it is also possible to characterize *fake Brownian motions* with Volterra representation of the form (2.2.2). In this case we do assume the kernels to be square integrable. (Cf. Lemma 2.2.6 below.)

Theorem 28. *A process $(X_t)_{t\geq 0}$ of the form (2.2.2) is a Brownian motion if and* *only if the Volterra kernel l is self reproducing, i.e., l satisfies*

$$l(t,s) = \int_0^s l(t,v)l(s,v)\,dv \tag{2.2.4}$$

for all t and for all $s \leq t$. In this case $\{X_s \,|\, s \leq t\}$ is independent of $\int_0^t l(t,u)\,dB_u$ for any $t > 0$.

Proof. Lemma 2.3 in Föllmer, Wu and Yor [FWY99] implies that (X_t) is a Brownian motion if and only if

$$\mathbb{E}\left[X_s \int_0^t l(t,u)\,dB_u\right] = 0 \quad \forall s \leq t, \tag{2.2.5}$$

i.e., if the Gaussian family $\{X_s \,|\, s \leq t\}$ is independent of $\int_0^t l(t,u)\,dB_u$. Since l is continuous and X is defined via (2.2.2) we have

$$\mathbb{E}\left[X_s \int_0^t l(t,u)\,dB_u\right] = \int_0^s l(t,u)\,du - \int_0^s \int_0^u l(t,v)l(u,v)\,dv\,du.$$

It follows that equation (2.2.5) is equivalent to

$$\int_0^s l(t,u)\,du = \int_0^s \int_0^u l(t,v)l(u,v)\,dv\,du \tag{2.2.6}$$

for all $s \leq t$ and subsequently is equivalent to equation (2.2.4) for all t and all $s \leq t$. $\square$

We give an explicit example of Brownian motions in terms of self reproducing Volterra kernels. Consider the special case where l is the product of two deterministic continuous functions a and b,

$$l(t,s) := a(t)b(s),$$

where a and b satisfy the following conditions.

(A) $a \in L^1[0,t]$ for all t and for all $t_0 > 0 :\ a(t) \not\equiv 0$ on (t_0,∞).
(B) $b \in L^2[0,t]$ for all $t > 0$, and also for all $t > 0$

$$\int_0^t \frac{|b(u)|}{\sqrt{\int_0^u b(v)^2\,dv}}\,du < \infty.$$

Corollary 2.2.4. *Let the process (X_t) be of the form*

$$X_t = B_t - \int_0^t a(u) \int_0^u b(v)\, dB_v\, du$$

with deterministic functions a and b satisfying conditions (A) and (B). Then the process X is a Brownian motion if and only if it is of the form

$$X_t = B_t - \int_0^t \frac{b(u)}{\int_0^u b(v)^2\, dv} \int_0^u b(r)\, dB_r\, du. \tag{2.2.7}$$

Proof. According to Theorem 28 it is sufficient to prove that a Volterra kernel l of the form $l(t,s) = a(t)b(s)$ satisfies condition (2.2.4) if and only if

$$a(t) = \frac{b(t)}{\int_0^t b(u)^2\, du}.$$

Substituting $l(t,s) = a(t)b(s)$ in condition (2.2.4) we find

$$a(t)b(s) = a(t)a(s) \int_0^s b(u)^2\, du,$$

and according to condition (A) we obtain the result. □

Example 2.2.5. Take $b(t) = t^m$ for $m > -\frac{1}{2}$. Then $a(t) = \frac{t^m}{\frac{t^{2m+1}}{2m+1}} = (2m+1)t^{-m-1}$ and X, given by

$$X_t = B_t - (2m+1) \int_0^t \int_0^u u^{-m-1} v^m\, dB_v\, du,$$

is a Brownian motion. For $m = 0$ we get, like in Remark 2.2.3, that

$$X_t := B_t - \int_0^t \frac{B_u}{u}\, du$$

is a Brownian motion.

Now we turn to the main result of this section.

Theorem 29. *There exist fake Brownian motions $(X_t)_{t\geq 0}$ with quadratic variation $\langle X \rangle_t = t$ which are continuous Gaussian semimartingales (but not Brownian motions).*

The theorem will be proved after the following

Lemma 2.2.6. *A process X defined by (2.2.2), i.e. of the form*

$$X_t := B_t - \int_0^t \int_0^u l(u,v)\, dB_v\, du$$

is a fake Brownian motion if and only if for all t

$$\int_0^t l(t,v)\, dv = \int_0^t \int_0^s l(t,v) l(s,v)\, dv\, ds. \tag{2.2.8}$$

Proof. Let X be defined as above. Since X is a centered Gaussian process, X is a fake Brownian motion if and only if

$$\mathbb{E}[X_t^2] = t. \tag{2.2.9}$$

Since $\langle X \rangle_t = t$ we get by Itō's formula

$$X_t^2 = 2 \int_0^t X_u\, dX_u + t.$$

Hence the condition on the variance reads

$$\begin{aligned} 0 = \mathbb{E}[X_t^2] - t = \mathbb{E}\left[2 \int_0^t X_u\, dX_u\right] = \mathbb{E}\left[\int_0^t X_u\, dX_u\right] = \\ = \mathbb{E}\left[\int_0^t X_u \int_0^u l(u,v)\, dB_v\, du\right]. \end{aligned}$$

The validity of the last line for all t is equivalent to

$$0 = \mathbb{E}\left[X_u \int_0^u l(u,v)\, dB_v\right] = \int_0^u l(u,v)\, dv - \int_0^u \int_0^r l(u,r) l(v,r)\, dr\, dv$$

for all u. □

Proof of Theorem 29. It remains to prove that there exist Volterra kernels which satisfy condition (2.2.8) but not (2.2.4). To this end we consider a kernel of the form

$$l(u,v) = \frac{1}{u} \varphi\left(\frac{v}{u}\right).$$

On the one hand, l satisfies (2.2.4) if and only if

$$\varphi(x) = \int_0^1 \varphi(zx) \varphi(z)\, dz, \tag{2.2.10}$$

on the other hand, l satisfies (2.2.8) if and only if

$$\int_0^1 \varphi(x)\,dx = \int_0^1 \int_0^1 \varphi(zx)\varphi(z)\,dz\,dx. \tag{2.2.11}$$

Now consider for instance $\varphi(x) = ce^{-ax}$ for some fixed a. Then (2.2.11) holds if and only if

$$c = \frac{(1-e^{-a})}{\int_0^a e^{-u}(1-e^{-u})\frac{du}{u}}.$$

But (2.2.10) is not satisfied for $c \neq 0$. □

Remark 2.2.7. We close this section by noting that Föllmer, Wu and Yor [FWY00] show also that, in the class of continuous semimartingales of the form

$$X_t = B_t + \int_0^t v_s\,ds,$$

fake Brownian motions can be characterized by a weak martingale property in analogy to Lévy's characterization of (proper) Brownian motion. (Note that processes of the above form have quadratic variation $\langle X\rangle_t = t$). In this setting, a continuous semimartingale X is called a weak martingale if, for all bounded Borel-measurable functions f and for all $t > 0$,

$$\mathbb{E}\left[\int_0^t f(X_s)\,dX_s\right] = 0.$$

Remark 2.2.8. So far we only considered semimartingales, but often it is desirable to get a proper martingale as mimicking process and accordingly many attempts were made to fit martingales to given marginals (cf. for instance the seminal paper of Madan and Yor [MY02]). The following section contains a first step in this direction for the particular case of Brownian motion.

2.3 Markov-Martingales & Marginals – discontinuous case

In this section we discuss two ways of mimicking Brownian motion via Markov martingales. Both approaches yield discontinuous processes and, as we will see in the next section, it is not at all trivial to find a continuous fake Brownian motion which is a martingale but not Brownian motion itself.

In [MY02] Madan and Yor discuss three approaches to construct a Markov martingale that matches given marginals: (i) a continuous martingale approach via an SDE similar to Dupire [Dup97]; (ii) an approach via a random time change; and (iii) a method using Skorohod embedding. In case of Brownian marginals however, the first two approaches reduce to a construction of Brownian motion itself (cf. [HK07], [MY02]). Only the method via Skorohod embedding yields something new, namely a one sided jump process.

The second approach of constructing a discontinuous Markov martingale with appropriate marginals is due to Hamza and Klebaner, [HK07]. The idea of their approach is to define a two step process and an associated transition function. Below we will present briefly the Skorohod approach due to Madan and Yor and spend the rest of the section to sketch the construction of Hamza and Klebaner. The discussion of the first two Madan/Yor approaches in case of Brownian motion is postponed to Section 2.4.

2.3.1 Skorohod Embedding

As already noted, the third approach of Madan and Yor uses a family of solutions to the Skorohod embedding, or Skorohod stopping problem to construct a martingale with given marginal densities $(g(t,y))_{t\geq 0}$. The Skorohod stopping problem (cf. [RY99], p.269f) consists in finding a finite stopping time T s.t. for a given probability measure μ on $\mathbb{R}$, μ is the law of a stopped Brownian motion, i.e., $B_T \sim \mu$.

Since a solution T of the Skorohod problem has to be finite, we have $\mathbb{E}[T] < \infty$, hence the stopped Brownian motion B^T is square integrable and $(B_t^T)^2 - T\wedge t$ is uniformly integrable. Consequently $\mathbb{E}[B^T] = 0$ and $\mathbb{E}[B_T^2] = \mathbb{E}[T]$. Therefore, a measure μ which admits a solution of the Skorohod problem necessarily has to be centered with finite second moment. As was proved by Skorohod this is already sufficient in order to find a stopping time T that solves the problem.

In the construction of a martingale matching $(g(t,y))_{t\geq 0}$ Madan/Yor use a solution to the Skorohod problem given by Azéma and Yor with the help of a barycentre function ψ. Let $\mu(dy)$ be a centered probability measure on the real line with $\int |y|\,\mu(dy) < \infty$, then the function

$$\psi(x) := \frac{\int_x^\infty y\,\mu(dy)}{\int_x^\infty \mu(dy)} \tag{2.3.1}$$

is a positive increasing function, $\psi(x) \geq x$ and $\lim_{x \to -\infty} \psi(x) = 0$. To construct the desired stopping time, one considers simultaneously the Brownian motion and its maximum to date $S_t := \sup_{0 \leq s \leq t} B_s$. The key tool to construct a family of stopping times $(T_t)_{t \geq 0}$ s.t. $(M_t)_{t \geq 0} := (B_{T_t})_{t \geq 0}$ is the following

Theorem 30. *(Cf. [RY99], p.272 for statement and proof.) Let μ be a centered probability measure, then the stopping time*

$$T_\mu := \inf\{t \geq 0 : S_t \geq \psi(B_t)\} \tag{2.3.2}$$

is a.s. finite, $B_{T_\mu} \sim \mu$ and

(i) B^{T_μ} is a U.I. martingale
(ii) $\mathbb{E}[\langle B, B \rangle_{T_\mu}] = \int x^2 \, \mu(dx)$.

Now, as implicitly indicated above, for a family of densities $(g(t,y))_{t \geq 0}$ one defines $\mu_t(dy) = g(t,y)dy$ and a parametrized barycentre function

$$\psi_t(x) = \psi(t,x) = \frac{\int_x^\infty y \, g(t,y) \, dy}{\int_x^\infty g(t,y) \, dy}. \tag{2.3.3}$$

Recall that, according to Theorem 24, there exists a martingale having the marginal distributions $\mu_t(dy) = g(t,y)dy$ if and only if these marginals increase in the *convex order*. Another order relation on the set of random variables is given via the *mean residual value*. A family of densities $(g(t,y))_{t \geq 0}$ is said to increase in the *mean residual value* if and only if the barycentre function $\psi_t(x)$ increases in t for each x. This order relation is stronger than the convex ordering (cf. [MY02], p.512-513), so, if we assume the family $(g(t,y))_{t \geq 0}$ to be increasing in the mean residual value, the random variables accordingly distributed increase in the convex order and we can apply Theorem 24 (due to Kellerer), i.e. there exists a martingale having these marginal distributions. Collecting all facts from above, we obtain:

Theorem 31. *(Cf. [MY02], p.513.) Let $(g(t,y))_{t \geq 0}$ be a family of zero mean densities on $\mathbb{R}$ having the property of increasing mean residual value, then there exists an increasing family of Brownian stopping times $(T_t)_{t \geq 0}$ such that*

(i) $(M_t) := (B_{T_t})$ is a martingale;
(ii) $(M_t)_{t \geq 0}$ is an inhomogeneous Markov process;
(iii) for each t the density of M_t is given by $g(t,y)$.

Proof. As already indicated, define T_t by

$$T_t := \inf\{s \geq 0 : S_s \geq \psi_t(B_s)\}$$

where $S_t = \sup_{0\leq s\leq t} B_s$. From Theorem 30 above we get that the law of M_t is $g(t,y)dy$ for each t, hence property (iii).
Since the function $\psi_t(x)$ is increasing in t by assumption, it follows that $s < t$ implies $T_s \leq T_t$ and since B^{T_t} is U.I. the Optional Sampling Theorem implies (i) because of

$$\mathbb{E}[B_{T_t} \,|\mathcal{F}_{T_s}] = B_{T_s}.$$

To prove the Markov property we note that for $s < t$

$$\begin{aligned}
T_t &= \inf\{u \,|\, S_u \geq \psi_t(B_u)\} \\
&= T_s + \inf\{v \,|\, S_{T_s+v} \geq \psi_t(B_{T_s+v})\} \\
&= T_s + \inf\left\{v \,|\, S_{T_s} \vee \left(\sup_{T_s\leq h\leq T_s+v} B_h\right) \geq \psi_t(B_v)\right\} \\
&= T_s + \inf\left\{v \,|\, S_{T_s} \vee \left(B_{T_s} + \sup_{0\leq u\leq v} \widetilde{B}_u\right) \geq \psi_t(B_{T_s} + \widetilde{B}_v)\right\}
\end{aligned}$$

where $\widetilde{B}_u = B_{T_s+u} - B_{T_s}$.
Now define

$$\widetilde{T}(z,b) = \inf\{v \,|\, z \vee (b + \widetilde{S}_v) \geq \psi_s(b + \widetilde{B}_v)\}$$

and observe that, for any test function $f(x)$, we get

$$\mathbb{E}^{B_{T_s}=b}\Big[f(B_{T_t}) \,|\, \mathcal{F}_{T_s}\Big] = \mathbb{E}\left[f\left(\widetilde{B}_{\widetilde{T}(\psi_s(b),b)}\right)\right].$$

I.e., $M_t = B_{T_t}$ is an inhomogeneous Markov process. $\square$

In the case of Brownian motion, i.e. Gaussian $\mathcal{N}(0,t)$ densities, the random variables B_t clearly are increasing in the convex order since they are the marginals of a martingale. Indeed they are also increasing in the mean residual value, i.e. the barycentre function is increasing in t for each x. This can be seen by differentiating $\psi_t(x)$. We have

$$\psi_t(x) = \frac{\sqrt{t}\exp(-x^2/2)}{\sqrt{2\pi}\left(1 - N\left(\frac{x}{\sqrt{t}}\right)\right)},$$

where N is the cumulative distribution function $N(y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} \exp(-x^2/2)\,dx$.

However, to see that this construction in principal yields a one sided jump process it is necessary to calculate the infinitesimal generator. (See [MY02], p.514ff. for the calculation, and p.525ff. for examples of martingales with various other marginals constructed via Skorohod embedding.)

2.3.2 The elementary approach

The construction of Hamza and Klebaner does not refer to advanced techniques such as the Skorohod embedding or the theory of SDEs, but constructs a transition function that satisfies the Chapman-Kolmogorov equations (cf. Definition 1.4.4) via a parametrized sum of independent random variables. Hence the existence of a (canonical) Markov process $(X_t)_{t\geq 0}$ follows from Theorem 8 above. The families of the thus constructed processes are continuous in probability, but not a.s., and admit predictable quadratic variation $\langle X, X\rangle_t \neq t$. (All the material in the rest of the section is taken from [HK07], hence we won't cite each proposition or definition separately.)

The construction is based on the trivial observation that, for $r \in [0,1]$ and standard normally distributed random variables Y and ξ,

$$Z = \sqrt{r}\,Y + \sqrt{1-r}\,\xi$$

is again $\mathcal{N}(0,1)$ distributed. This immediately leads to

Proposition 2.3.1. *Let (R, Y, ξ) be a triple of independent random variables such that*

(i) $R \in [0,1]$ *and*
(ii) $Y \sim \mathcal{N}(0,1)$, $\xi \sim \mathcal{N}(0,1)$.

Then the random variable $Z := \sqrt{R}\,Y + \sqrt{1-R}\,\xi$ *is also standard normal. The pair* (Y, Z) *is bivariate Gaussian if and only if* R *is non-random.*

Proof. The conditional joint distribution of (Y, Z) given R is bivariate normal with zero means and covariance matrix

$$\begin{pmatrix} 1 & \sqrt{R} \\ \sqrt{R} & 1 \end{pmatrix}.$$

The marginals of a bivariate normal pair have distributions that do not depend on their correlation, thus the unconditional distribution of Z will remain standard normal. However $\mathbb{E}[Z^2 \,|\, Y = 0] = \mathbb{E}[1 - R]$ but $\mathbb{E}[Z^4 \,|\, Y = 0] = 3\mathbb{E}[(1 - R)^2]$. If they were bivariate normal, $\mathbb{E}[(1 - R)^2] = (\mathbb{E}[1 - R])^2$, and $R \equiv \text{const}$ a.s. □

Now let $\alpha > 0$, assume that R takes values in $[0, 1]$, let $Y \sim \mathcal{N}(0, \alpha^2)$, $\xi \sim \mathcal{N}(0, 1)$ and let R, Y and ξ be independent. Then

$$Z = \sigma \left(\sqrt{R} Y + \alpha \sqrt{1 - R}\, \xi \right) \sim \mathcal{N}(0, \sigma^2 \alpha^2).$$

The two-step process (Y, Z) is a martingale if and only if

$$Y = \mathbb{E}[Z \,|\, Y] = \mathbb{E}[\sigma(\sqrt{R} Y + \alpha \sqrt{1 - R}\, \xi) \,|\, Y] = \sigma Y \, \mathbb{E}[\sqrt{R}],$$
$$\text{i.e., } \mathbb{E}[\sqrt{R}] = \frac{1}{\sigma}. \tag{2.3.4}$$

Furthermore, in order to define a continuous time Markov process by means of such a two step process, we need

(i) The conditional distribution of the two step process which serves as a (basic) transition function.
(ii) A family of $(0, 1]$-valued random variables $(R_{s,t})_{0<s\leq t}$ which
 a) depend on (s, t) only through $\sqrt{t/s} =: \sigma$,
 b) satisfy the moment of order $1/2$, $\mathbb{E}[\sqrt{R_{s,t}}] = 1/\sigma = \sqrt{s/t}$, and
 c) whose distributions, denoted by G_σ, $R_{s,t} \sim G_{\sqrt{t/s}}(dr) = G_\sigma(dr)$, form a log-convolution semigroup.
(iii) A family of standard Gaussian random variables $(\xi_{s,t})$.

We proceed by assuming (ii) and (iii); the (not yet parametrized) conditional distribution of Z given Y is

$$F_{Z \,|\, Y=y}(dz) = \mathbb{P}(R = 1)\delta_{\sigma y}(dz) + \mathbb{E}\left[\varphi\left(\sigma \sqrt{R} y, \sigma^2\alpha^2(1 - R), z\right) \mathbb{1}_{R<1}\right] dz, \tag{2.3.5}$$

where δ_x is the Dirac measure concentrated on x and $\varphi(\mu, \sigma^2, \cdot)$ is the Gaussian density with mean μ and variance σ^2.

Accordingly we define the inhomogeneous Markov process $(X_t)_{t\geq 0}$, such that for times $t > s$ the r.v. X_t is given a.s. in terms of the triple $(R_{s,t}, X_s, \xi_{s,t})$,

$$X_t = \sqrt{\frac{t}{s}} \left(\sqrt{R_{s,t}}\, X_s + \sqrt{s}\, \sqrt{1 - R_{s,t}}\, \xi_{s,t} \right). \tag{2.3.6}$$

In Proposition 2.3.5 we will see that the parametrized version of (2.3.5), i.e. the t.f. for (2.3.6), satisfies the Chapman-Kolmogorov equations. Before we can prove that, we have to clarify the notion of a log-convolution semigroup, and its relation to convolution semigroups (cf. Definition 1.5.2), which, as we saw in Lemma 1.5.3, give rise to Feller transition functions.

Remark 2.3.2. We follow the notation of Hamza and Klebaner and use the abbreviations

$$\sigma = \sqrt{t/s}, \quad \alpha = \sqrt{s}, \quad \tau = \sqrt{u/t}, \text{ hence } R_{s,t} = R_\sigma, \quad R_{t,u} = R_\tau.$$

It should be noted that almost the hardest part of the construction is to keep track of parameters and indices.

Definition 2.3.3. A family of distributions $(G_\sigma)_{\sigma \ge 1}$ on $(0, \infty)$ is a log-convolution semigroup if $G_1 = \delta_1$ and if the distribution of the product of any two independent random variables with distributions G_τ and G_σ is $G_{\sigma\tau}$.

Proposition 2.3.4. *Let $(G_\sigma)_{\sigma \ge 1}$ be a log-convolution semigroup on $(0, 1]$ and, for $\sigma \ge 1$, let R_σ be a random variable with distribution G_σ. If F_s, $s \ge 0$, denotes the distribution of $V_s = -\log R_{e^s}$ then $(F_s)_{s \ge 0}$ is a convolution semigroup.*
Conversely, let $(F_s)_{s \ge 0}$ be a convolution semigroup and, for $s \ge 0$, let V_s be a random variable with distribution F_s. If G_σ, $\sigma \ge 1$, denotes the distribution of $R_\sigma = e^{-V_{\log \sigma}}$ then $(G_\sigma)_{\sigma \ge 1}$ is a log-convolution semigroup.

Proof. Straightforward. □

Proposition 2.3.5. *Define for $x \in \mathbb{R}$, $s > 0$ and $t = \sigma^2 s \ge s$, $P_{s,t}(x, dy)$ by*

$$\begin{aligned} P_{0,t} &= \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y-x)^2}{2t} \right) dy, \\ P_{s,t} &= \gamma(\sigma)\delta_{\sigma x}(dy) + \left(\int_{(0,1)} \frac{1}{\sqrt{2\pi t}\, \sqrt{1-r}} \exp\left(-\frac{(y - \sigma\sqrt{r}x)^2}{2t(1-r)} \right) G_\sigma(dr) \right) dy = \\ &= \gamma(\sigma)\delta_{\sigma x}(dy) + \mathbb{E}\left[\varphi\left(\sigma\sqrt{R_\sigma}\, x, \sigma^2\alpha^2(1-R_\sigma), y \right) \mathbb{1}_{R<1} \right] dy, \end{aligned}$$

where $R_\sigma \sim G_\sigma$ *and* $\gamma(\sigma) = G_\sigma(\{1\}) = \mathbb{P}(R_\sigma = 1)$.
If $(G_\sigma)_{\sigma\geq 1}$ *is a log-convolution semigroup then for any* $u > t > s > 0$ *and any* x

$$P_{s,u}(x, dz) = \int P_{s,t}(x, dy)\, P_{t,u}(y, dz) \tag{2.3.7}$$

and, for any $u > t > 0$,

$$P_{0,u}(0, dz) = \int P_{0,t}(0, dy)\, P_{t,u}(y, dz).$$

Proof. We present the (short) proof of the consistency of the a.s. representation (2.3.6), given in the arXiv preprint of the paper (cf. [HK06]); to see the full (and rather lengthy) direct computation of the Chapman-Kolmogorov equations, cf. [HK07], p.4.

$$\begin{aligned}
X_u &= \tau\left(\sqrt{R_{t,u}}\, X_t + \sigma\alpha\sqrt{1 - R_{t,u}}\, \xi_{t,u}\right) \\
&= \tau\left(\sqrt{R_{t,u}}\, \sigma\left(\sqrt{R_{s,t}}\, X_s + \alpha\sqrt{1 - R_{s,t}}\, \xi_{s,t}\right) + \sigma\alpha\sqrt{1 - R_{t,u}}\, \xi_{t,u}\right) \\
&= \sigma\tau\left(\sqrt{R_{s,t}R_{t,u}}\, X_s + \alpha\left(\sqrt{(1 - R_{s,t})R_{t,u}}\, \xi_{s,t} + \sqrt{1 - R_{t,u}}\, \xi_{t,u}\right)\right) \\
&= \sqrt{\frac{u}{s}}\left(\sqrt{R_{s,u}}\, X_s + \sqrt{s}\sqrt{1 - R_{s,u}}\, \xi_{s,u}\right)
\end{aligned}$$

where $R_{s,u} = R_{s,t}R_{t,u}$ and

$$\xi_{s,u} = \left(\frac{\sqrt{(1 - R_{s,t})R_{t,u}}}{\sqrt{1 - R_{s,u}}}\, \mathbb{1}_{R_{s,u}<1} + \mathbb{1}_{R_{s,u}=1}\right) \xi_{s,t} + \frac{\sqrt{1 - R_{t,u}}}{\sqrt{1 - R_{s,u}}}\, \mathbb{1}_{R_{s,u}<1}\, \xi_{t,u}.$$

(The unconditional distribution of $\xi_{s,u}$ as well its conditional distribution given $R_{s,t}$ and $R_{t,u}$ are standard Gaussian, which implies that $\xi_{s,u}$ is independent of $R_{s,u}$). □

Theorem 8 above guarantuees the existence of a Markov process $(X_t)_{t\geq 0}$ which is distributed according to the transition function. What is still missing, however, is a condition that ensures the martingale property of $(X_t)_{t\geq 0}$. Recall that the necessary and sufficient condition of the defining two-step process to be a martingale, for $\sigma \geq 1$, reads

$$\mathbb{E}\left[\sqrt{R_{s,t}}\right] = \mathbb{E}\left[\sqrt{R_\sigma}\right] = \frac{1}{\sigma}. \tag{2.3.8}$$

Recall also that, for $R_\sigma \sim G_\sigma$, $(G_\sigma)_{\sigma\geq 1}$ is a log-convolution semigroup if and only if $(F_s)_{s\geq 0}$, where F_s, $s \geq 0$, denotes the distribution of $V_s = -\log R_{e^s}$, is a convolution semigroup (cf. Proposition 2.3.4). This convolution semigroup now defines a subordinator, i.e. an increasing (Lévy) process with stationary independent increments, and we can apply the Lévy-Khintchine formula in the Laplace-version (cf. Theorem 12).

Proposition 2.3.6. *Assume that the family $(G_\sigma)_{\sigma\geq 1}$ is a log-convolution semigroup on $(0,1]$ and let $(R_\sigma)_{\sigma\geq 1}$ be independent r.v.s accordingly distributed. Then, for $\sigma \geq 1$, the positive random variable $U_\sigma = -\log R_\sigma$ is infinitely divisible and the moment generating function of U_σ,*

$$L_\sigma(\lambda) = \mathbb{E}\left[e^{\lambda \log R_\sigma}\right] = \mathbb{E}\left[(R_\sigma)^\lambda\right]$$

is given by

$$\log L_\sigma(\lambda) = -\left(\beta\lambda + \int_0^\infty (1-e^{-\lambda x})\,\nu(dx)\right)\log\sigma, \tag{2.3.9}$$

where the Lévy measure ν satisfies $\nu(\{0\}) = 0$ and $\int_0^\infty (1\wedge x)\,\nu(dx) < \infty$. Conversely, any function L_σ of the form (2.3.9) is the moment of order λ of a log-convolution semigroup $(G_\sigma)_{\sigma\geq 1}$.

According to Theorem 12 we denote the Laplace exponent for the log-convolution semigroup $(G_\sigma)_{\sigma\geq 1}$ by

$$\phi(\lambda) = \beta\lambda + \int_0^\infty (1-e^{-\lambda x})\,\nu(dx). \tag{2.3.10}$$

The condition on $(X_t)_{t\geq 0}$ to be a martingale can, taking $\lambda = 1/2$, now be stated as follows

$$\frac{1}{\sigma} = \mathbb{E}\left[\sqrt{R_\sigma}\right] = L_\sigma\left(\frac{1}{2}\right) = \sigma^{-\phi\left(\frac{1}{2}\right)} = \frac{1}{\sigma^{\phi\left(\frac{1}{2}\right)}},$$

or, equivalently,

$$\phi\left(\frac{1}{2}\right) = 1. \tag{2.3.11}$$

Collecting all results the main theorem reads as follows.

Theorem 32. *(Hamza, Klebaner, 07) Let the family $(G_\sigma)_{\sigma\geq 1}$ form a log-convolution semi-group with Laplace exponent*

$$\phi(\lambda) = \beta\lambda + \int_0^\infty (1-e^{-\lambda x})\,\nu(dx).$$

Assume that $\phi(1/2) = 1$. Then there exists a fake BM Markov martingale $(X_t)_{t\geq 0}$

(w.r.t. its natural filtration) starting at zero with transition function given by Proposition 2.3.5, $X_t \sim \mathcal{N}(0,t)$, and which has the representation

$$X_t = \sqrt{\frac{t}{s}}\left(\sqrt{R_{s,t}}\,X_s + \sqrt{s}\,\sqrt{1-R_{s,t}}\,\xi_{s,t}\right)\ a.s., 0 < s < t.$$

Proof. Start with a function $\phi(\lambda)$ of the form (2.3.10) which satisfies (2.3.11) and construct, with the help of the Lévy-Khintchine formula, the log-convolution semigroup $(G_\sigma)_{\sigma\geq 1}$. The transition function $P_{s,t}(x,dy)$ is given by Proposition 2.3.5 and Theorem 8 ensures the existence of a canonical version of $(X_t)_{t\geq 0}$, having all desired properties. □

Example 2.3.7. (Due to Olle Häggström, cf. [Alb08], p.685.) Define

$$X_t = \sqrt{t}\,B_1(-1)^{N_t},$$

where $(N_t)_{t\geq 0}$ is an inhomogeneous Poisson process, independent of $(B_t)_{t\geq 0}$ with intensity $\lambda(r) = 1/(4r)$ and $\mathcal{F}_t = \sigma(B_1) \vee \sigma(N_s; s \leq t)$. Clearly, M is not continuous and coincides with Brownian motion in the one dimensional marginals. To show, that X is a martingale, one uses the usual rules of measurability and independence concerning conditional expectation and the following fact which is due to basic properties of Poisson processes.

$$\begin{aligned}\mathbb{P}(N_t - N_s \text{ odd}) &= \frac{1}{2}\left(1-\exp\left(-2\int_s^t \lambda(r)\,dr\right)\right) = \frac{1}{2}\left(1-\exp\left(-2\int_s^t \frac{1}{4r}\,dr\right)\right) = \\ &= \frac{1}{2}\left(1-\sqrt{\frac{s}{t}}\right).\end{aligned}$$

Hence we get

$$\mathbb{E}[X_t\,|\,\mathcal{F}_s] = \sqrt{t}B_1(-1)^{N_s}\mathbb{E}\left[(-1)^{N_t-N_s}\right] = X_s\sqrt{\frac{t}{s}}\,(1-2\,\mathbb{P}(N_t-N_s \text{ odd}) = X_s$$

We conclude the discussion of the mimicking results of Hamza and Klebaner by noting several properties of the thus constructed process $(X_t)_{t\geq 0}$. For the explicit constructions and the properties of the two constructed classes of processes ($\gamma(\sigma) = G_\sigma(\{1\}) = 0$, and $\gamma(\sigma) > 0$), cf. [HK07], p.9ff.

Theorem 33. *The process $(X_t)_{t\geq 0}$ is continuous in probability, i.e.*

$$\forall c > 0\colon \lim_{s\to t}\mathbb{P}(|X_t - X_s| > c) = 0.$$

Theorem 34. *Set $\delta = \frac{\phi(1)}{2}$. Then the (predictable) quadratic variation of X_t is given by*

$$\langle X, X \rangle_t = \delta t + (1-\delta) \int_0^t \frac{X_s^2}{s} \, ds.$$

Theorem 35. *The process $(X_t)_{t \geq 0}$ is quasi-left continuous. It is continuous if and only if $G_\sigma \equiv \delta_{\sigma^{-2}}$, i.e., $R_{s,t} \equiv s/t$, in which case $(X_t)_{t \geq 0}$ is a standard Brownian motion.*

2.4 Martingale & Marginals – continuous case

From Theorem 35 we see that if the Hamza/Klebaner construction were supposed to yield a continuous Markov martingale with Brownian marginals, it reduces to a construction of Brownian motion itself. This is also the case for the two continuous fitting approaches of Madan/Yor in [MY02] which we briefly summarize before turning to the second and main part of this section where we present two constructions of continuous fake Brownian motions which are martingales but *not* Markov. The first one is due to J.M.P. Albin in [Alb08], who was the first one to construct such a martingale; the second one comes from K. Oleszkiewicz, [Ole08].

2.4.1 Collapsing examples

The first continuous approach in [MY02], as noted in the foregoing section, looks for a martingale defined via an SDE of the form

$$X_t = \int_0^t \sigma(s, X_s) \, dB_s. \tag{2.4.1}$$

The forward transition densities $p(t, y)$ of (2.4.1) satisfy the Kolmogorov forward equation

$$\frac{\partial}{\partial t} p(t, y) = \frac{1}{2} \frac{\partial^2}{\partial y^2} \left(\sigma(y, t)^2 p(t, y) \right), \tag{2.4.2}$$

and, as we saw in Proposition 0.3.1 of the introduction, knowing the function

$$C(t, k) = \mathbb{E}[(X_t - k)^+] = \int_k^\infty (y - k) p(t, y) \, dy.$$

amounts to knowing the densities $p(t,y)$ and so we are able to obtain the diffusion coefficient
$$\sigma(t,k)^2 = \frac{2C_t}{C_{kk}}.$$
However, in the case of Brownian motion, if we plug in
$$p(t,y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{y^2}{2t}}$$
in (2.4.2), we get that $\sigma^2 \equiv 1$ and (X_t) is just a standard Brownian motion.

The random time change approach in [MY02] now looks for an inhomogeneous increasing Markov process $(L_t)_{t\geq 0}$ with independent increments which is independent of a Brownian motion $(B_t)_{t\geq 0}$. Then the sought martingale is $(B_t)_{t\geq 0}$, subordinated by this process, s.t.
$$X_t := B_{L_t}$$
matches the given family of marginals $(g(t,y))_{t\geq 0}$. To this end one relates the characteristic function of X_t and the Laplace transform of L_t and notes that
$$\mathbb{E}[e^{iuX_t}] = \mathbb{E}\left[\exp\left(-\frac{u^2}{2}L_t\right)\right] = \int_{-\infty}^{\infty} e^{iuy} g(t,y)\,dy. \tag{2.4.3}$$
However, in the case of Brownian marginals, i.e., $g(t,y) = p(t,y)$ (like above), equation (2.4.3) gives, using the independence of (B_t) and (L_t),
$$\mathbb{E}[e^{iuX_t}] = \mathbb{E}\left[\exp\left(-\frac{u^2}{2}L_t\right)\right] = \exp\left(-\frac{1}{2}u^2 t\right), \tag{2.4.4}$$
hence $L_t = t$ and $X_t = B_t$ itself.

Remark 2.4.1. In fact, it has to be noted that the above procedure, i.e. to run Brownian motion at a different speed to get a martingale with different marginals, is just the reverse direction of a well known theorem of stochastic calculus which states that *every* continuous martingale is just a time-changed Brownian motion. In particular we see that the quest for a continuous martingale with Brownian marginals can not lead us very far astray from Brownian motion itself.

Theorem 36. *(Dambis, Dubins-Schwarz; cf. [RY99], p.181) Let $(\mathcal{F}_t)_{t\geq 0}$ be a right continuous filtration, let M be a continuous $(\mathcal{F}_t,\mathbb{P})$-local martingale vanishing at* 0 *and such that* $\langle M, M\rangle_\infty = \infty$. *Set*

$$T_t = \inf\{s \,|\, \langle M, M\rangle_s > t\},$$

then $(B_t) = (M_{T_t})$ *is an* $(\mathcal{F}_{T_t})$*-Brownian motion and* $M_t = B_{\langle M,M\rangle_t}$.

2.4.2 Continuous Fake Brownian motions

The first construction, as mentioned, is due to J.M.P. Albin and defines a continuous fake Brownian motion as the product of two independent, weak solutions $(X^{(1)})_{t\geq 0}$, $(X^{(2)})_{t\geq 0}$ of the time homogeneous SDE

$$dX_t = \frac{1}{2X_t}\, dB_t, \quad X_0 = 0, \tag{2.4.5}$$

and a r.v. Y with probability densitiy function given by

$$f_Y(y) = \frac{4\left(\Gamma\left(\frac{3}{4}\right)\right)^2}{2\pi^{3/2}\sqrt{1-(y/\sqrt{2})^4}}, \quad \text{for } y \in (0, \sqrt{2}). \tag{2.4.6}$$

We state the theorem and give just a sketch of proof, since it is not only involved, but also uses many earlier results of Albin which are just cited in [Alb08]. We note that the proof of the equality in the one dimensional marginals relies heavily on the scaling property, resp. self-similarity, of the densities of the mimicking process resp. the Brownian motion. This we already encountered in the introduction, where we used $B_t \overset{(law)}{=} \sqrt{t}\, B_1$ to construct a mimicking process X. Recall that Gaussian densities scale as follows

$$p(t; 0, y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{y^2}{2t}} = \frac{1}{\sqrt{t}}\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y}{\sqrt{t}}\right)^2}\right) = \frac{1}{\sqrt{t}}\, p\left(1; 0, \frac{y}{\sqrt{t}}\right). \tag{2.4.7}$$

Hence, to check the equality in the marginals it is only necessary to check this at unit time, if the densities scale.

Theorem 37. *(Albin, 2008) Let* $(X_t^{(1)})_{t\geq 0}$, $(X_t^{(2)})_{t\geq 0}$ *be independent weak solutions of equation* (2.4.5) *and let* Y *be distributed according to* (2.4.6)*, then*

$$M_t := X_t^{(1)}\, X_t^{(2)}\, Y, \quad t \geq 0,$$

is a continuous martingale such that $M_t \overset{(law)}{=} B_t$ *for* $t \geq 0$*, but it is not Brownian motion.*

Proof. (Sketch) The proof proceeds, roughly, in four steps. First it is proved, that weak solutions of the SDE (2.4.5) exist and the transition densities $p(t;0,y)$ of $X_t^{(i)}$ are determined. In the second step it is shown, that the product $X^{(1)}X^{(2)}Y$ indeed is a martingale. Third of all, using quadratic variation, it is checked that (M_t) is not Brownian motion and in the fourth step equality in the marginals is proved via Mellin transforms.

(i) The diffusion coefficient $\sigma(y) = \frac{1}{2y}$ is nonzero for all y and

$$\int_{-\varepsilon}^{\varepsilon} \frac{dy}{\sigma(x+y)^2} = \int_{-\varepsilon}^{\varepsilon} 4y^2\, dy = \frac{4}{3}\left((x+\varepsilon)^3 - (x-\varepsilon)^3\right) < \infty \text{ for } x \in \mathbb{R}, \varepsilon > 0.$$

According to the Engelbert-Schmidt theory (cf. Section 1.7) Equation (2.4.5) has a weak solution X. To find the transition density

$$p(t;x,y) = \frac{\mathbb{P}(X_{t+s} \in (y, y+dy)\,|X_s = x)}{dy}$$

one considers the transformed process $Y = X^4$. By Itō's formula we get a Cox Ingersoll Ross process

$$\begin{aligned} dY_t = d(X_t^4) &= 4X_t^3\, dX_t + \frac{1}{2} 12X^2\, d\langle X, X\rangle_t = 4\frac{X_t^3}{2X_t}\, dB_t + \frac{1}{2}\frac{12X^2}{4X^2}\, dt \\ &= 2\sqrt{Y_t}\, dB_t + \frac{3}{2}\, dt, \quad Y_0 = 0, \end{aligned} \tag{2.4.8}$$

the transition density of which can be explicitly computed (cf. [Alb08]). From the transition density of the CIR process Y, however, we can recover the transition density of the (original) solution X.

$$p(t;0,y) = \frac{2^{1/4}y^2}{\Gamma\left(\frac{3}{4}\right)t^{3/4}} \exp\left(-\frac{y^4}{2t}\right), \quad \text{for } y \in \mathbb{R},\ t > 0. \tag{2.4.9}$$

(ii) Since we know that the above solution X to (2.4.5) is just a continuous *local* martingale, we have to show that

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_t|^n\right] = \mathbb{E}\left[\sup_{t\in[0,T]} |Y_t|^{n/4}\right] < \infty \quad \text{for } T > 0, \tag{2.4.10}$$

in order to get a continuous martingale. Combining various results of Albin (cf. [Alb08], p.684) one obtains the estimate

$$\mathbb{P}\left(\sup_{t\in[0,T]} |Y_t|^{n/4} > u\right) \sim 2\,\mathbb{P}\left(Y_T > u^{4/n}\right) = \int_{u^{4/n}}^{\infty} \frac{1}{(2T)^{3/4} y^{1/4} \Gamma\left(\frac{3}{4}\right)} \exp\left(-\frac{y}{2T}\right) dy,$$

where $\sim$ refers to asymptotic equality as $u \to \infty$. So we see that (2.4.10) holds for $n \geq 0$ and $(X_t^{(i)})_{t\geq 0}$, $i = 1, 2$, are true martingales with respect to their natural filtrations $(\mathcal{F}_t^{X^{(1)}})$ and $(\mathcal{F}_t^{X^{(2)}})$. Since $(X^{(1)})$ and $(X^{(2)})$ are independent, the product $(X^{(1)}X^{(2)})$ is a martingale w.r.t. $(\mathcal{F}_t^{X^{(1)}} \vee \mathcal{F}_t^{X^{(2)}})$ and $M_t = X^{(1)}\, X^{(2)}\, Y$ is a martingale w.r.t. $(\mathcal{F}_t^{X^{(1)}} \vee \mathcal{F}_t^{X^{(2)}} \vee \sigma(Y))$.

(iii) If (M_t) were a Browian motion it would have quadratic variation $\langle M\rangle_t = t$. However, Y does not depend on t, hence we get

$$t = \langle M\rangle_t = Y^2 \langle X^{(1)} X^{(2)}\rangle_t,$$

and

$$\mathbb{E}[\langle X^{(1)} X^{(2)}\rangle_t] = t\,\mathbb{E}[1/Y^2] = \infty,$$

as can be seen by computing the integral $\int_{(0,\sqrt{2})} \frac{1}{y^2} f_Y(y)\, dy$, where $f_Y(y)$ is the density (2.4.6). The Burkholder-Davis-Gundy and Cauchy-Schwarz inequalities imply that

$$\mathbb{E}[\langle X^{(1)} X^{(2)}\rangle_t] \leq K\,\mathbb{E}\left[\sup_{s\in[0,t]} |X_s^{(1)}|^2 |X_s^{(2)}|^2\right] \leq K\,\mathbb{E}\left[\sup_{s\in[0,t]} |X_s^{(1)}|^4\right] < \infty,$$

for some $K > 0$. So (M_t) cannot be Brownian motion.

(iv) Finally, to show that $M_t \overset{(law)}{=} B_t$ for all t, we use the above mentioned scaling property of the Gaussian densities, i.e. $B_t \overset{(law)}{=} t^{1/2}\, B_1$ and the self-similarity of the densities $p(t; 0, y)$ of $X^{(1)}$ and $X^{(2)}$ (cf. (2.4.9)) which scale with $t^{1/4}$.

$$p(t; 0, y) = t^{1/4}\, p\left(1; 0, \frac{y}{t^{1/4}}\right), \text{ i.e., } X^{(i)} \overset{(law)}{=} t^{1/4}\, X_1^{(i)},\ i = 1, 2.$$

Therefore it suffices to show, that $X_1^{(1)} X_1^{(2)} Y \overset{(law)}{=} B_1$. By symmetry this follows if the Mellin transform $\widehat{M_1}$ of $|X_1^{(1)}||X_1^{(2)}|Y$,

$$\widehat{M_1} = \mathbb{E}\left[\left(|X_1^{(1)}||X_1^{(2)}|Y\right)^{s-1}\right] = \mathbb{E}\left[|X_1^{(1)}|^{s-1}\right]^2 \mathbb{E}\left[Y^{s-1}\right],$$

agrees with that of $|B_1|$, denoted by $\widehat{B_1}$.

$$\widehat{B_1} = \mathbb{E}\left[|B_1|^{s-1}\right] = \int_0^\infty y^{s-1}\frac{2}{2\pi}\exp\left(-\frac{y^2}{2}\right)\,dy = \frac{2^{s/2}\Gamma(s/2)}{\sqrt{2\pi}}, \text{ for } s \geq 1.$$

According to the transition density (2.4.9), $|X_1^{(1)}|$ and $|X_1^{(2)}|$ have common Mellin transform

$$\widehat{X_1} = \int_0^\infty y^{s-1}\frac{2^{5/4}y^2}{\Gamma\left(\frac{3}{4}\right)}\exp\left(-\frac{y^4}{2}\right)\,dy = \frac{2^{s/4}\Gamma\left(\frac{s+2}{4}\right)}{2^{1/4}\Gamma\left(\frac{3}{4}\right)}, \quad \text{for } s \geq 1.$$

So, to get $\widehat{B_1} \overset{!}{=} \widehat{M_1} = \widehat{X_1}^2\widehat{Y}$, it is enough to prove that Y has Mellin transform

$$\widehat{Y} = \frac{\widehat{B_1}}{\widehat{X_1}^2} = \frac{\Gamma\left(\frac{3}{4}\right)^2 2^{s/2}\Gamma\left(s/4\right)}{2\pi\,\Gamma\left(\frac{s+2}{4}\right)}.$$

However, this is, as can be computed via the density $f_Y(y)$, precisely the Mellin transform of Y. □

Remark 2.4.2. A continuous fake Brownian motion (M_t) that is a martingale cannot have independent increments. If $M_{t+s} - M_s$ and $M_s \overset{(law)}{=} B_s$ are independent, then by Cramer's theorem (cf. [Cra36]) all summands of the $\mathcal{N}(0, t+s)$ distributed sum $M_{t+s} = (M_{t+s} - M_s) + M_s$ must be normally distributed, i.e. the increments of M,

$$M_{t+s} - M_s \overset{(law)}{=} B_{t+s} - B_s \overset{(law)}{=} B_t \overset{(law)}{=} M_t,$$

are even stationary and we get that M is a continuous Gaussian process with stationary independent increments, i.e. Brownian motion itself.

Now that we have seen a technically rather involved construction, we turn, like in the discontinuous case, to a more basic approach and present the construction of Oleszkiewicz in [Ole08] which was the second reply to the question raised by Hamza/Klebaner.

Theorem 38. *(Oleszkiewicz, 2008) There exists a continuos fake Brownian motion which is a martingale but not Brownian motion itself.*

Proof. We proceed in three steps: (i) we define a continuous fake BM $(X_t^{(a)})$ which is a martinale for times $t > e^{-a}$ and extend the definition to all $t > 0$. (ii) we prove the a.s. continuity of the paths for $t > 0$ and (iii) the so constructed process is extended to $t = 0$ and the continuity in 0 is shown. The martingale property of the final process $(X_t)_{t\geq 0}$ is shown via the definition of conditional expectation on a multiplicative system.

(i) We start with some notation. Throughout the proof, let N_1 and N_2 be independent $\mathcal{N}(0,1)$ distributed random variables, let $(B_t)_{t\geq 0}$ be standard Brownian motion and assume that N_1, N_2 and $(B_t)_{t\geq 0}$ are independent.

Given $a \geq 0$ define the filtration $\mathcal{F}_t^{(a)} = \sigma(N_1, N_2, (B_s)_{0\leq s\leq a+\log t})$ for $t \geq e^{-a}$. Then the process

$$X_t^{(a)} = \sqrt{t}\left(N_1 \cos B_{a+\log t} + N_2 \sin B_{a+\log t}\right), \text{ for } t \geq e^{-a}$$

is a continuous martingale w.r.t. the filtration $(\mathcal{F}_t^{(a)})_{t\geq e^{-a}}$. Furthermore, $X_t^{(a)} \sim \mathcal{N}(0,t)$ for every $t \geq e^{-a}$ which can easily be seen by just computing expectation and variance. It is also easy to check that $X_e^{(a)} - X_1^{(a)}$ is not Gaussian, i.e., $(X_t^{(a)})_{t\geq e^{-a}}$ cannot be a Gaussian process.

Another way to see the difference to Brownian motion, i.e. that $(X_t^{(a)})_{t\geq e^{-a}}$ cannot be extended to BM, is to note that

$$\limsup_{t\to\infty} \frac{|X_t^{(a)}|}{\sqrt{t}} \leq |N_1| + |N_2| < \infty \quad a.s.,$$

while by the Law of the Iterated Logarithm (cf. Theorem 27) we have

$$\limsup_{t\to\infty} \frac{|B_t|}{\sqrt{t}} = \infty. \quad a.s.$$

Now, for $0 < b < a$, let $(V_s)_{s\geq 0}$ be given by $V_s = B_{a-b+s} - B_{a-b}$. From Proposition 2.1.5-(i) we know that $(V_s)_{s\geq 0}$ is a standard Brownian motion. Furthermore $(V_s)_{s\geq 0}$, B_{a-b}, N_1 and N_2 are independent. We define

$$N_1' = N_1 \cos B_{a-b} + N_2 \sin B_{a-b},$$

and

$$N_2' = N_2 \cos B_{a-b} - N_1 \sin B_{a-b},$$

and note, again by direct calculation, that $N_1' \overset{(law)}{=} N_2' \sim \mathcal{N}(0,1)$. For $t \geq e^{-b} \geq e^{-a}$ we have

$$\begin{aligned}
(X_t^{(a)})_{t\geq e^{-b}} &= \Big(\sqrt{t}\,(N_1 \cos B_{a+\log t} + N_2 \sin B_{a+\log t})\Big)_{t\geq e^{-b}} \\
&= \Big(\sqrt{t}\,(N_1 \cos(B_{a-b} + V_{b+\log t}) + N_2 \sin(B_{a-b} + V_{b+\log t}))\Big)_{t\geq e^{-b}} \\
&= \Big(\sqrt{t}\,(N_1 \cos B_{a-b} + N_2 \sin B_{a-b}) \cos V_{b+\log t} + \\
&\quad + (N_2 \cos B_{a-b} - N_1 \sin B_{a-b}) \sin V_{b+\log t}\Big)_{t\geq e^{-b}} \\
&= \Big(\sqrt{t}\,(N_1' \cos V_{b+\log t} + N_2' \sin V_{b+\log t})\Big)_{t\geq e^{-b}}. \qquad (2.4.11)
\end{aligned}$$

The process (2.4.11) has the same distribution as $(X_t^{(b)})_{t\geq e^{-b}}$ since

$$(N_1', N_2', (V_s)_{s\geq 0}) \overset{(law)}{=} (N_1, N_2, (B_s)_{s\geq 0})$$

and the above construction can be extended to all $t > 0$. To this end, let $e^{-a} \leq t_1 < \cdots < t_n$, and set

$$\mu_{t_1,\ldots,t_n} = \mathbb{P}(X_{t_1}^{(a)} \in A_1, \ldots, X_{t_k}^{(a)} \in A_k), \quad k \in \mathbb{N},\ A_i \in \mathcal{B}(\mathbb{R}^n).$$

In view of the above, $\mu_{t_1,\ldots,t_n}$ does not depend on a and it is clear that the family

$$\mathcal{P}^f_{X^{(a)}} := \{\mu_{t_1,\ldots,t_n} \,|\, n \in \mathbb{N},\ 0 < t_1 \cdots < t_n\}$$

is consistent, since for any two of them some $a > 0$ may be chosen such that both of them are finite dimensional distributions of the process $(X_t^{(a)})_{t\geq e^{-a}}$. By the Kolmogorov extension (or consistency) theorem there exists a stochastic process $(\widetilde{X}_t)_{t>0}$ such that the measures $\mu_{t_1,\ldots,t_n}$ are its f.d.d.s, hence $(\widetilde{X}_t)_{t>e^{-a}}$ has the same finite dimensional distributions as $(X_t^{(a)})_{t\geq e^{-a}}$ for every $a > 0$. What we have to show, however, is that, for $t > 0$, there exist a continuous modification $(X_t)_{t>0}$ of $(\widetilde{X}_t)_{t>0}$. Therefore we define

$$X_t = \widetilde{X}_t \quad \text{for } t \in \mathbb{Q}^+,$$

and

$$X_t = \lim_{s\to t; s\in\mathbb{Q}^+} X_s \quad \text{for } t \in (0,\infty) \cap \mathbb{Q}^+.$$

For $k \in \mathbb{N}$ and $t \in (e^{-k}, \infty)\backslash\mathbb{Q}^+$ now the random events

$$\{\text{path of } (X_s^{(k)})_{s\in[e^{-k},k]\cap\mathbb{Q}^+} \text{ is uniformly continuous }\},$$

and

$$\{\lim_{s\to t; s\in[e^{-k},\infty)\cap\mathbb{Q}^+} X_s^{(k)} = X_t^{(k)}\}$$

depend only on countably many random variables and belong to the cylindrical σ-field of the process $(X_s^{(k)})_{s\geq e^{-k}}$. Thus

$$\mathbb{P}\left((\widetilde{X}_s)_{s\in[e^{-k},k]\cap\mathbb{Q}^+} \text{ is uniformly continuous}\right) = \\ = \mathbb{P}\left((X_s^{(k)})_{s\in[e^{-k},k]\cap\mathbb{Q}^+} \text{ is uniformly continuous}\right) = 1$$

and

$$\mathbb{P}\left(X_t = \widetilde{X}_t\right) = \mathbb{P}\left(\lim_{s\to t; s\in[e^{-k},\infty)\cap\mathbb{Q}^+} X_s^{(k)} = X_t^{(k)}\right) = 1.$$

Since there are only countably many $k \in \mathbb{N}$, $(X_t)_{t>0}$ has a.s. continuous paths and is a modification of $(\widetilde{X}_t)_{t>0}$.

(ii) To get a.s. continuity of the paths at $t = 0$, we set $X_0 = 0$ and note that for $C > 0$ and $a > 0$ one has

$$\mathbb{P}\left(\sup_{t\geq e^{-a}} \frac{|X_t^{(a)}|}{\sqrt{t}} > C\right) \leq \mathbb{P}\left(|N_1| + |N_2| > C\right) \leq 2\mathbb{P}\left(|N_1| > \frac{C}{2}\right) \leq 4e^{-\frac{C^2}{8}}.$$

Hence, for $a > n + 1$,

$$\mathbb{P}\left(\exists_{t\in[e^{-n-1},e^{-n}]}|X_t| > C_n e^{-\frac{n}{2}}\right) \leq \mathbb{P}\left(\sup_{t\in[e^{-n-1},e^{-n}]} \frac{X_t^{(a)}}{\sqrt{t}} > C_n\right) \leq 4e^{-\frac{C_n^2}{8}}.$$

Now set $C_n = e^{\varepsilon n/2}$ for $\varepsilon \in (0, 1)$. Then the Borel-Cantelli Lemma yields

$$\frac{|X_t|}{t^{\frac{1}{2}-\varepsilon}} \xrightarrow{t\to 0^+} 0 \quad \text{a.s. for every } \varepsilon > 0.$$

(iii) In order to obtain a martingale for $t \in \mathbb{R}^+$ we also have to extend the filtration to 0 and therefore define, as usual, $\mathcal{F}_t^X = \sigma(X_s; 0 \leq s \leq t)$. To check that, for $0 < t < T$, we get $\mathbb{E}[X_T \,|\, \mathcal{F}_t^X] = X_t$ we consider a multiplicative system $\mathcal{A}_t$ consisting of random events of the form $\{X_{s_1} \in B_1, \ldots, X_{s_k} \in B_k\}$ for $k \in \mathbb{N}$, $0 < s_1, \ldots, s_k \leq t$ and $B_1, \ldots, B_k \in \mathcal{B}(\mathbb{R})$. Clearly $\sigma(\mathcal{A}_t) = \mathcal{F}_t^X$, so it suffices to check that $\mathbb{E}[X_T \mathbb{1}_A] = \mathbb{E}[X_t \mathbb{1}_A]$ for every $A \in \mathcal{A}_t$. Given $a > 0$ with $e^{-a} = \min_{i\leq k} s_i$ we define

$$\tilde{A} = \{X_{s_1}^{(a)} \in B_1, \ldots, X_{s_k}^{(a)} \in B_k\},$$

so that

$$\mathbb{E}[X_T \mathbb{1}_A] = \mathbb{E}[X_T^{(a)} \mathbb{1}_{\tilde{A}}] = \mathbb{E}[X_t^{(a)} \mathbb{1}_{\tilde{A}}] = \mathbb{E}[X_t \mathbb{1}_A].$$

This completes the proof. □

2.5 Markov-Martingales & Marginals – continuous case

To conclude the chapter on mimicking Brownian motion we will discuss why the attempts to mimick Brownian motion by a continuous Markov martingale yielded nothing except Brownian motion itself (cf. Section 2.4.1).

We explicitly consider the special case where the mimicking Markov martingale Y is adapted to the filtration generated by Brownian motion. In this case the (weak) Markov property of the mimicking martingale Y suffices to determine it as Brownian motion.

The case where the mimicking process is a continuous strong Markov martingale with Brownian marginals is a corollary to a theorem of Lowther (cf. Theorem 44). It shows that every continuous strong Markov martingale is uniquely determined via its one dimensional marginals. In Section 3.1.1 we will discuss the general result and its background in some detail.

Proposition 2.5.1. *Let $(B_t)_{t\geq 0}$ be a standard Brownian Motion and let $(Y_t)_{t\geq 0}$ be a continuous Markov martingale adapted to $(\mathcal{F}_t) = (\sigma(B_s; s \leq t))$. If, for all $t \in \mathbb{R}^+$, $B_t \overset{law}{=} Y_t$ then the process $(Y_t)_{t\geq 0}$ is a (standard) Brownian Motion.*

Proof. $(Y_t)_{t\geq 0}$ is a martingale adapted to the Brownian filtration, hence by the Martingale Representation Theorem there exists an adapted process $\Gamma(t, \omega)$ such that

$$Y_t = \mathbb{E}[Y_0] + \int_0^t \Gamma(s, \omega)\, dB_s.$$

$\mathbb{E}[Y_0] = 0$, since $Y_0 \sim \mathcal{N}(0, 0)$. We will show that the adapted process $\Gamma(t, \omega)^2$ is constant and equal to one.

From the Itō-formula we know on the one hand that for every $u \in \mathcal{C}^{1,2}(\mathbb{R}^+ \times \mathbb{R})$

$$du(t, B_t) = u_t(t, B_t)\, dt + u_x(t, B_t)\, dB_t + \frac{1}{2} u_{xx}(t, B_t)\, d\langle B_t \rangle$$

and, on the other hand,

$$\begin{aligned} du(t, Y_t) &= u_t(t, Y_t)\, dt + u_x(t, Y_t)\, dY_t + \frac{1}{2} u_{xx}(t, Y_t)\, d\langle Y_t \rangle = \\ &= u_t(t, Y_t)\, dt + u_x(t, Y_t)\, \Gamma(t, \omega)\, dB_t + u_{xx}(t, Y_t)\, \Gamma(t, \omega)^2\, dt. \end{aligned}$$

Now, if we consider the PDE $u_t(t,x)+u_{xx}(t,x) = 0$ for $t \in [0,T]$, given the boundary condition $u(T,x) = g(x) \quad \forall x \in \mathbb{R}, \quad g \in \mathcal{C}^2(\mathbb{R})$, then the Feynman-Kac formula, or e.g. [Øks98, Exercise 9.3], tells us, that the solution to this equation (in terms of Brownian Motion) is given by

$$u(t, x) = \mathbb{E}^x g(B_{T-t}).$$

Now, since the one-dimensional marginals of B_t and Y_t coincide, we get

$$\begin{aligned} u(t, x) &= \mathbb{E}^x g(B_{T-t}) = \int_{-\infty}^{\infty} g(y)\, d\mathbb{P}^x(B_{T-t} \in dy) \overset{B_s \overset{law}{=} Y_s}{=} \int_{-\infty}^{\infty} g(y)\, d\mathbb{P}^x(Y_{T-t} \in dy) \\ &= \mathbb{E}^x g(Y_{T-t}). \end{aligned}$$

If we plug in Y_t for the dummy x we obtain from the Markov property

$$u(t, Y_t) = \mathbb{E}^{Y_t} g(Y_{T-t}) = \mathbb{E}^x[g(Y_T)|\mathcal{F}_t]$$

and consequently $u(t, Y_t)$ is a martingale. So we conclude (from the Itō-formula for (Y_t) above), that

$$u_t(t, Y_t)\, dt + u_{xx}(t, Y_t)\, \Gamma(t, \omega)^2\, dt = 0 \quad \forall t \leq T.$$

But $u(t,x) = \mathbb{E}^x g(B_{T-t}) = \mathbb{E}^x g(Y_{T-t})$ is a solution for $u_t(t,x) + u_{xx}(t,x) = 0$ for all $x \in \mathbb{R}, \quad t \leq T$. In particular we have for (t, Y_t)

$$u_t(t, Y_t) + \frac{1}{2} u_{xx}(t, Y_t) = 0$$

as well as

$$u_t(t, Y_t) + \frac{1}{2} u_{xx}(t, Y_t) \Gamma(t, \omega)^2 = 0.$$

I.e., we get

$$\Gamma(t, \omega)^2 \equiv 1.$$

To see that $Y_t = \int_0^t \Gamma(t, \omega)\, dB_s$ indeed is just Brownian motion we calculate the quadratic variation

$$\langle Y_t, Y_t \rangle = \langle \int_0^t \Gamma(t, \omega)\, dB_s, \int_0^t \Gamma(t, \omega)\, dB_s \rangle = \int_0^t \Gamma(s, \omega)^2\, ds = \int_0^t 1\, ds = t.$$

Since $(Y_t)_{t \geq 0}$ is a continuous martingale, by Lévy's characterization it is Brownian motion. □

Remark 2.5.2. The method of proof presented above does not work for dimensions strictly greater than one.

Let $(B_t)_{t \geq 0}$ a d-dimensional Brownian Motion and let $(Y_t)_{t \geq 0}$ be a d-dimensional SMM, then, as is known, the solution of the SDE $dX_t = dB_t$ is associated to the PDE $u_t + \frac{1}{2}\Delta u = 0, \quad t < T$, and the solution is, like in the one-dimensional case, given by $u(t, x) = \mathbb{E}^x g(B_{T-t})$ for some boundary condition $u(T, x) = g(x)$. Likewise, the Martingale Representation Theorem gives us the existence of a d-dimensional previsible process ξ, s.t.

$$M(t) = M_0 + \sum_{i=1}^{d} \int_0^t \xi_i\, dB_i$$

for ξ_i in L^2.

But, if one compares, like in the one-dimensional case, the coefficients of the second partial derivatives, one gets

$$\sum_{i=1}^{d} u^i_{xx}(t, Y_t) = \sum_{i=1}^{d} u^i_{xx}(t, Y_t) \xi_i^2$$

which is not at all unique. E.g. in the case $d = 2$ fix t and suppose $u^1_{xx}(t, Y_t) = u^2_{xx}(t, Y_t) = 1$, then

$$2 = \xi_1^2 + \xi_2^2$$

and the only condition for $(\xi_1(t, \omega), \xi_2(t, \omega))$ is, to be an element of the circle of radius 2. In the one-dimensional case, the argument works, because of the rather simple structure of the unit-sphere in dimension 1. $S^0 = \{-1, 1\}$.

Theorem 39. *Let $(X_t)_{t\geq 0}$ be a continuous Strong Markov martingale having the same marginals as linear Brownian motion, then X is a version of $(B_t)_{t\geq 0}$.*

Proof. Apply Theorem 44. □

2.6 Summary

In this chapter we encountered two discontinuous and three continuous constructions of processes with Brownian marginals which are different from Brownian motion. From the point of view of characterization the continuous approaches are of particular interest. As mentioned, it was an open problem until 2008 whether there exists a continuous martingale with Brownian marginals different from BM. The strong Markov martingale case yields nothing except Brownian motion itself which is also clear since 2008. The case of a Markov martingale adapted to the Brownian filtration was discussed in Proposition 2.5.1. However, to the knowledge of the author it is not known whether the (weak) Markov property suffices to characterize a continuous martingale with Brownian marginals as Brownian motion.

3 Mimicking Itō-processes

In this final chapter we enlarge the class of processes to be mimicked from Brownian motion to martingales and Itō-processes. In particular we are interested in so called *Markovian projections* which are Markov processes having the same 1-d marginals as the original process. The reason why such projections are of considerable interest was already discussed in the introduction. Furthermore, we will discuss the correspondence between the class of martingale-marginals and the set of martingale measures on the canonical probability space and we will identify a class of processes (i.e. martingale measures) for which the Markovian projection is a well defined and continuous function.

Section 3.1 summarizes several results of Lowther on fitting marginals to real valued martingales. The main result of the first section, Theorem 44, states that in the class of continuous strong Markov martingales the one-dimensional marginal distributions suffice to determine the structure of the entire process. The second important result of the first section, Theorem 46, yields the continuity of the Markovian projection as a function from the set of *weakly continuous martingales* to the set of *almost continuous martingale diffusions*. (The set of almost continuous diffusions is slightly larger than the set of continuous strong Markov processes.)

In Section 3.2 we discuss a, by now, celebrated result of Gyöngy (Theorem 47) which represents the second main result of this chapter. Given any k-dimensional Itō-process the coefficients of which satisfy a few mild assumptions, Theorem 47 ensures the existence of a certain Markov process having the same one-dimensional marginals. In particular this Markov process is given as a solution of an SDE the coefficients of which have an explicit representation in terms of the coefficients of the (original) Itō-process.

3.1 Martingales – One Dimension

Martingales play a crucial role in the theory of stochastic processes, so it is desirable to know how the distribution of a martingale is related to its one dimensional marginals.

Consequently, the subject of this section is not to go into technical details of fitting/mimicking w.r.t. certain families of distributions, but to present some general results. We consider results on so called *almost continuous diffusion (ACD)* martingales which also hold for continuous martingales.

The first part of this section presents a *generalized backward equation*, as Lowther calls it. The underlying idea of this construction consists in a combination of the Kolmogorov backward and forward equation. The generalized equation can be extended to cover not only continuous diffusions but also jump diffusions and/or quasimartingales. In particular, the generalized equation serves as a *martingale condition* for functionals of martingales where the classical results do not hold because of lack of smootheness. The generalized equation also implies Theorem 44, the main result of [Low08a] and of this section.

In the second part of the section we present other related results of Lowther, taken from [Low08b]. In particular Lowther showed that if one requires the marginals not only to be increasing in the convex order with constant mean, but also to be weakly continuous, then these marginals can always be fitted in a unique way by a martingale within the class of almost continuous diffusions. Furthermore, he showed that the mapping from the set of marginals to the set of martingale measures on the canonical probability space is continuous.

This section follows [Low08a] and [Low08b] very closely, so we won't give detailed references.

3.1.1 The generalized backward equation

We briefly recall the results of Section 1.7 where we established the link between SDEs and PDEs. In particular, recall Proposition 1.7.3 and the Kolmogorov backward equation (Theorem 19).

Let $(X_t)_{t\geq 0}$ be a solution to a (1-d) stochastic differential equation of the form

$$dX_t = \sigma(t, X_t)\, dB_t + b(t, X_t)\, dt, \quad X_0 = x. \tag{3.1.1}$$

Prop. 1.7.3 yields that $u(t, X_t)$ is a local martingale for $u \in \mathcal{C}^{1,2}([0,\infty) \times \mathbb{R})$ if

$$\frac{\partial}{\partial} u_t(t,x) + \frac{1}{2}\sigma(t,x)^2 \frac{\partial^2}{\partial x^2} u(t,x) + b(t,x)\frac{\partial}{\partial x} u(t,x) = 0. \tag{3.1.2}$$

If σ and b are sufficiently well behaved so that for any $T > 0$ and any smooth bounded function $g(x)$ there exists a bounded solution to (3.1.2) for $t \leq T$ satisfying the boundary condition $f(T,x) = g(x)$, then $u(t, X_t)$ is a proper martingale and

$$u(t, X_t) = \mathbb{E}[g(X_T) \,|\, \mathcal{F}_t], \tag{3.1.3}$$

hence $(X_t)_{t\geq 0}$ is Markov. The other direction, however, to define u via equation (3.1.3) and to show that it is twice differentiable is not possible in general.

The idea how to define a martingale condition for $u(t, X_t)$ where u is not continuously differentiable and where even no drift or diffusion coefficients might exist, is to remove the dependence of the martingale condition on an SDE of the form (3.1.1). The generalized backward equation related to X is then not stated in terms of the coefficients σ and b, but in terms of the marginal distributions and a drift measure (and possibly a jump component). We will only consider the case of continuous martingales, and leave the case of quasimartingales to the interested readers of [Low08a]. In the following pages we stick to the definition of the strong Markov property given in Def. 1.4.15. Note that this definition does not refer to a transition function.

A key element in defining the generalized equation is the function $C\colon \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ of Proposition 0.3.1, which uniquely determines the one-dimensional marginals of a process X,

$$C(t,x) = \mathbb{E}[(X_t - x)^+]. \tag{3.1.4}$$

Furthermore, if $(X_t)_{t\geq 0}$ is a martingale and additionally the densities $p(t,x)$ of $(X_t)_{t\geq 0}$ are smooth, we obtain from the Kolmogorov forward / Fokker-Planck equation

$$\frac{\partial}{\partial t} C(t,x) = \frac{1}{2}\sigma(t,x)^2 \frac{\partial^2}{\partial x^2} C(t,x). \tag{3.1.5}$$

For smooth $u\colon \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ we may combine this equation with the backward equation (3.1.2), substituting σ. We obtain a martingale condition for $u(t, X_t)$ different from (3.1.2) and without any reference to the SDE (3.1.1). I.e., $u(t, X_t)$ is a martingale if

$$\frac{\partial u}{\partial t}\frac{\partial^2 C}{\partial x^2} + \frac{\partial C}{\partial t}\frac{\partial^2 u}{\partial x^2} = 0. \tag{3.1.6}$$

To extend the above condition to non-differentiable functions u and C we smooth the expression by multiplying with a twice differentiable function θ having compact support in $(0, \infty) \times \mathbb{R}$ and integrate the equation. We then set

$$\mu_{[u,C]}(\theta) = \int\int \theta \left(\frac{\partial u}{\partial t}\frac{\partial^2 C}{\partial x^2} + \frac{\partial C}{\partial t}\frac{\partial^2 u}{\partial x^2}\right) dt\, dx. \tag{3.1.7}$$

Observe that $\mu_{[u,C]}(\theta)$ is a linear function of θ and that it is symmetric in u and C. We integrate by parts to obtain

$$\mu_{[u,C]}(\theta) = \int\int \left(\frac{\partial u}{\partial x}\frac{\partial C}{\partial x}\frac{\partial \theta}{\partial t} - \frac{\partial \theta}{\partial x}\frac{\partial u}{\partial x}\frac{\partial C}{\partial t} - \frac{\partial C}{\partial x}\frac{\partial \theta}{\partial x}\frac{\partial u}{\partial t}\right) dt\, dx. \tag{3.1.8}$$

The martingale condition for $u(t, X_t)$ now reads $\mu_{[u,C]} = 0$ and the functions u and C only have to be assumed differentiable. Furthermore, differentiability w.r.t. t is removed with the help of Lebesgue-Stieltjes integrals. For every $x \in \mathbb{R}$ such that $u(t, x)$ is right continuous with locally finite variation in t, the Lebesgue-Stieltjes integral $\int \cdot\, d_t u(t, x)$ is (locally) a finite signed measure satisfying

$$\int_{t_0}^{t_1} d_t u(t, x) = u(t_1, x) - u(t_0, x).$$

Regarding the differentiability in x we note that the definition of $C(t, x)$ implies that it is convex in x, hence the partial derivative w.r.t. x exists almost everywhere. Since condition (3.1.8) is stated in terms of integrals this is sufficient for the right hand side of (3.1.8) to be well defined. Definition 3.1.2 below states the appropriate conditions on u and C.

Remark 3.1.1. One of the problems in showing that $\mu_{[u,C]} = 0$ in general is a valid and sufficient martingale condition for $u(t, X_t)$ lies in the fact that, for non-smooth functions u and C, a priori one does not even know whether $u(t, X_t)$ is a semimartingale. To circumvent this difficulty Lowther makes use of the theory of *Dirichlet processes*. These are processes which can be decomposed into a local martingale and a zero quadratic variation term. See [Low08a, Sections 7, 8] for details. However, we won't prove but only will *use* the condition $\mu_{[u,C]} = 0$ to show Theorem 44.

First, we have to extend the martingale condition $\mu_{[u,C]} = 0$ to an appropriate class of non-differentiable functions. Let $\int_0^T |d_t f(t,x)|$ denote the variation of f with respect to t over the interval $[0,T]$.

Definition 3.1.2. Denote by $\mathcal{D}$ the set of functions $f\colon \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ such that

(i) f is Lipschitz in x and cadlag in t,
(ii) for every $K_0 < K_1 \in \mathbb{R}$ and $T \in \mathbb{R}^+$

$$\int_{K_0}^{K_1} \int_0^T |d_t f(t,x)|\, dx < \infty,$$

(iii) the left and right derivatives of $f(t,x)$ w.r.t. x exist everywhere.

The set $\mathcal{D}_K$ consists by definition of those functions $f \in \mathcal{D}$ which have compact support in $(0,\infty) \times \mathbb{R}$.

Note that if $f(t,x)$ is convex and Lipschitz in x and right continuous and monotone in t, then $f \in \mathcal{D}$. In particular, if C is defined as in (3.1.4) it is in $\mathcal{D}$ whenever $(X_t)_{t\geq 0}$ is a right continuous martingale. Furthermore, if $f \in \mathcal{D}$ then $f(t,x)$ will have locally finite variation in t for almost every x and the integral $\iint \cdot\, d_t f(t,x)\, dx$ is well defined and locally yields a finite signed measure.

Given a deterministic function $f\colon \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ let f^- denote the left limit in t,

$$f^-(t,x) = f(t-,x) := \begin{cases} \lim_{s \nearrow t} f(s,x), & \text{if } t > 0,\\ f(0,x), & \text{if } t = 0. \end{cases}$$

In order to obtain a valid martingale condition $\mu_{[u,C]} = 0$, as noted, the partial derivatives of u and C w.r.t. x have to exist almost everywhere. This holds for the functions in $\mathcal{D}$ which allows to extend the definition of $\mu_{[f,C]}(\theta)$ to $f, C \in \mathcal{D}$ and $\theta \in \mathcal{D}_K$.

Lemma 3.1.3. *(cf. [Low08a], p.6) If f, $g \in \mathcal{D}$, then $f(t,x)$ and $f^-(t,x)$ are differentiable in x almost everywhere w.r.t. the measure $\iint \cdot\, |d_t g(t,x)|\, dx$.*

Proof. By [Low08c, Lemma 3.3] $f(t,x)$ is differentiable in x almost everywhere w.r.t. the measure $\iint \cdot\, |d_t g(t,x)|\, dx$. It is only necessary to extend this to $f^-(t,x)$. Since $f(t,x)$ now is jointly continuous and cadlag in t, there can only be countably many times t at which $f^- \neq f$. So by countable additivity of the measure $\iint \cdot\, |d_t g(t,x)|\, dx$ we can restrict to fixed times $T > 0$. As $f^-(T,x)$ is Lipschitz in x, Lebesgue's theorem says that it will be differentiable almost everywhere w.r.t. the Lebesgue-measure. So,

$$\begin{aligned}\int\int \mathbb{1}_{\{t=T\}}\mathbb{1}_{\{f^-(t,x)\text{ is not differentiable in }x\}}\,|d_t g(t,x)|\,dx = \\ = \int \mathbb{1}_{\{f^-(T,x)\text{ is not differentiable in }x\}}\,|g(T,x)-g(T-,x)|\,dx = 0.\end{aligned}$$

□

Definition 3.1.4. For every $f,\, g \in \mathcal{D}$ define the linear map $\mu_{[f,g]}\colon \mathcal{D}_K \to \mathbb{R}$.

$$\mu_{[f,g]}(\theta) = \int\int f_x g_x\, d_t\theta\, dx - \int\int \theta_x^- f_x^-\, d_t g\, dx - \int\int g_x^- \theta_x^-\, d_t f\, dx. \tag{3.1.9}$$

The subscript x in f_x denotes the partial derivative w.r.t. x. If f and g are twice differentiable, the above definition coincides with (3.1.8).

The proof that $\mu_{[f,C]} = 0$ indeed can be used as a martingale condition for $f(t, X_t)$ consists roughly of two steps. First one defines the drift measure μ_f^X of a process X and then one shows that in the case of continuous strong Markov martingales it coincides with $\mu_{[f,C]}$. Hence, $\mu_{[f,C]}(\theta) = \mu_f^X(\theta) = 0$ for all $\theta \in \mathcal{D}_K$ in fact is a sufficient condition for $f(t, X_t)$ to be a martingale. (Necessity is obvious.)

Definition 3.1.5. Let X be a real valued and adapted cadlag process. We denote by $\mathcal{D}(X)$ the set of functions $f \in \mathcal{D}$ such that the following decomposition exists.

$$f(t, X_t) = M_t + A_t, \tag{3.1.10}$$

where (M_t) is a cadlag local martingale and (A_t) is a cadlag previsible process of locally finite variation with $A_0 = 0$ and such that $\int \mathbb{1}_{\{(s,X_s-)\in S\}}\, dA_s$ has integrable variation for every bounded Borel subset $S \subset \mathbb{R}^+ \times \mathbb{R}$. We define the local signed measure μ_f^X by

$$\mu_f^X(\theta) = \mathbb{E}\left[\int \theta(t, X_t-)\, dA_t\right],$$

for bounded measurable $\theta\colon \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ with bounded support.

Remark 3.1.6. μ_f^X is well defined since the above decomposition is unique if it exists. This follows by the standard argument that the difference of two such decompositions is a previsible local martingale with locally finite variation, hence constant.

In order to use the drift measure as a martingale condition we need two further results and the notion of *quasi left continuity*.

Definition 3.1.7. A process X is said to be quasi left continuous if $X_{\tau-} = X_\tau$ a.s. for all previsible stopping times $\tau > 0$.

Theorem 40. *Let X be a cadlag and quasi left continuous strong Markov process. Let $f \in \mathcal{D}(X)$ and let $f(t, X_t) = M_t + A_t$ be the above decomposition. Then, for every nonnegative and measurable $\theta \colon \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$,*

$$|\mu_f^X|(\theta) = \mathbb{E}\left[\int \theta(t, X_{t-})\,|dA_t|\right],$$

where $|\mu|$ denotes the variation of a local signed measure μ.

$$|\mu|(\theta) = \sup_{|g| \leq 1} \mu(g\theta).$$

The supremum is taken over all bounded and measurable g with $|g| \leq 1$.

Proof. Cf. [Low08a], p.28ff. □

From this result we obtain a first condition. If the drift measure $\mu_f^X = 0$, then A is constant, hence $f(t, X_t)$ is a local martingale. The next theorem relates the drift measure μ_f^X and $\mu_{[f,C]}$.

Theorem 41. *Let X be a continuous and strong Markov martingale and let $C \in \mathcal{D}$ be defined via $C(t, x) = \mathbb{E}[(X_t - x)^+]$. If $f \in \mathcal{D}$, then TFAE:*

(i) $f \in \mathcal{D}(X)$.
(ii) There exists a local signed measure μ such that $\mu_{[f,C]}(\theta) = \mu(\theta^-)$ for all $\theta \in \mathcal{D}_K$.

Furthermore, if these conditions hold then $\mu = \mu_f^X$.

Proof. Cf. [Low08a, Section 8] □

Now we prove that $\mu_{[f,C]} = 0$ indeed is a valid martingale condition in the class of strong Markov processes.

Theorem 42. *Let X be a continuous and strong Markov martingale and let $f \in \mathcal{D}$. Then $f(t, X_t)$ is a martingale if and only if $\mu_{[f,C]}(\theta) = 0$ for all $\theta \in \mathcal{D}_K$.*

Proof. First suppose that $f(t, X_t)$ is a martingale. Then the finite variation process in the semimartingale decomposition is zero. I.e., $A = 0$ in decomposition (3.1.10) and it follows that $f \in \mathcal{D}(X)$ and $\mu_f^X = 0$ for all $\theta \in \mathcal{D}_K$. Theorem 41 above yields $\mu_{[f,C]}(\theta) = \mu_f^X(\theta^-) = 0$.
Conversely, suppose that $\mu_{[f,C]} = 0$. Taking $\mu = 0$ in Theorem 41 above shows that $f \in \mathcal{D}(X)$ and $\mu_f^X = 0$. Now we write $f(t, X_t) = M_t + A_t$, which is decomposition (3.1.10), and we conclude from Theorem 40 that A has zero variation, i.e. $A = 0$ and $f(t, X_t) = M_t$ is a local martingale. As X is a martingale and $f(t, x)$ is Lipschitz in x, it follows that for every $t > 0$, $f(\tau, X_\tau)$ is UI over all stopping times $\tau \leq t$, hence $f(t, X_t)$ is a proper martingale. □

The last ingredient for the proof of the main result of this section is based on two further theorems of Lowther, [Low08d].

Theorem 43. *Let X be a continuous and strong Markov martingale, let $T \geq 0$ and let $g\colon \mathbb{R} \to \mathbb{R}$ be convex and Lipschitz. Then there exists a function $f\colon \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ s.t.*

$$f(t, X_t) = \mathbb{E}[g(X_T) \,|\, \mathcal{F}_t] \quad \forall t \leq T. \tag{3.1.11}$$

Moreover f is convex and Lipschitz in x, right continuous and decreasing in t, in particular, $f \in \mathcal{D}$.

Proof. The statement follows from [Low08d, Theorem 1.5 and 1.6]. □

Theorem 44. *(Lowther 2008, cf. [Low08a, Theorem 1.2].) Let $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ be continuous and strong Markov martingales such that $X_t \overset{(law)}{=} Y_t$ for all $t \in \mathbb{R}^+$. Then X and Y have the same joint distribution, i.e. they are versions of each other.*

Proof. First, we note that for any $T > 0$ the stopped processes X^T and Y^T are continuous and strong Markov martingales. Define $C \in \mathcal{D}$ now via the stopped process X^T, $C(t, x) = \mathbb{E}[(X_t^T - x)^+]$. Choose any convex and Lipschitz $g\colon \mathbb{R} \to \mathbb{R}$. Then choose $f \in \mathcal{D}$ such that equality (3.1.11) above is satisfied for every $t \leq T$. By definition $f(t, X_t^T)$ is a martingale, so Theorem 42 gives $\mu_{[f,C]} = 0$.

Since $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ coincide in the one dimensional marginals, the function C is also given by $C(t, x) = \mathbb{E}[(Y_t^T - x)^+]$ and we conclude via the reverse direction of Theorem 42 that $f(t, Y_t)$ also is a martingale and equation (3.1.11) is satisfied with Y instead of X. This shows that $(X_t, X_T) \overset{(law)}{=} (Y_t, Y_T)$ for each $t < T$, i.e. they are equal in all two-dimensional marginals. Furthermore, X and Y also have the same initial distribution and are Markov, which implies that X and Y are equal in all f.d.d.s. □

3.1.2 Fitting Marginals and Martingales

In this section we present the results of [Low08b] which yield that in a certain class of strong Markov processes a weakly continuous family (μ_t) of probability measures increasing in the convex order can be fitted continuously to a martingale. We won't give (detailed) proofs, but sketch the line of argument and present the concepts and the classes of processes and functions involved in the construction.

As (canonical) probability space we use in this section the Skorohod space, i.e. the space of cadlag real valued processes $\mathcal{D}[0,\infty) = \{\omega\colon [0,\infty) \to \mathbb{R} \,|\, \omega(t) \text{ is cadlag}\}$, where $X_t(\omega) = \omega(t)$ denotes the coordinate process with usual natural filtration $(\mathcal{F}_t)_{t\geq 0}$ and $\mathcal{F} = \mathcal{F}_\infty$. The aim of the following is to find a martingale measure $\mathbb{P}$ on $(\mathcal{D}[0,\infty), \mathcal{F})$ which matches a given family (μ_t) of probabiliy measures with the above mentioned properties. Weak continuity in this context means that if $t_n \to t$ then $\mu_{t_n}(f) \to \mu_t(f)$ for every continuous bounded f on $\mathbb{R}$.

The choice of cadlag processes instead of continuous ones is motivated by the fact, as Lowther notes, that there exist (simple examples of) marginal distributions which cannot be matched by any continuous process. For instance if $\mathbb{P}(0 < X_t < 1) = 0$ for all $t \geq 0$ and $\mathbb{P}(X_t \leq 0)$ decreases in t, then there must be a positive probability that X jumps from below 0 to above 1.

Definition 3.1.8. Let X be a real valued stochastic process. Then

(i) X is *almost continuous* if it is cadlag, continuous in probability and if, given two independent cadlag processes Y, Z each with the same distribution as X, we have for every $s < t \in \mathbb{R}^+$

$$\mathbb{P}\left(Y_s < Z_s, Y_t > Z_t \text{ and } Y_u \neq Z_u \, \forall u \in (s,t)\right) = 0$$

(ii) X is an *almost continuous diffusion* if it is *strong Markov* and *almost continuous.*

Remark 3.1.9. Condition (i) is equivalent to saying that $Y - Z$ cannot change sign without passing through zero, which is clearly true for continuous processes by the intermediate value theorem.

As in the previous section we use the function C to determine the family (μ_t). To stress that a particular process X has marginals consistent with some C, we write $C(t,x) = \mathbb{E}[(X_t - x)^+]$. The distribution functions can be recovered from $\mu_t((-\infty, x]) = 1 + \frac{\partial}{\partial x} C(t, x+)$. Accordingly the following space of functions will contain the families of marginals under consideration.

Definition 3.1.10. Let CP be the set of functions $C\colon \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ such that

(i) C is convex in x and increasing in t.
(ii) $C(t,x) \to 0$ as $x \to \infty$ for every $t \in \mathbb{R}^+$.
(iii) There exists a real number c such that $C(t,x) + x \to c$ as $x \to -\infty$ for every $t \in \mathbb{R}^+$.
(iv) C is continuous in t.

Remark 3.1.11. Conditions (i)-(iii) ensure that a certain function $C\colon \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ corresponds to a family of martingale marginals. The measures μ_t increase in the convex order if and only if C increases in t; condition (iii) is equivalent to requiring a constant mean. Condition (ii) is obvious from the definition of C. Property (iv), continuity in t, corresponds to assuming the marginals μ_t being weakly continuous in t which is the only assumption in addition to martingale marginals.

In order to state the main result we clarify some notions and notation. First, we denote by $\mathcal{M}(\mathcal{D})$ the set of probability measures on $(\mathcal{D}[0,\infty), \mathcal{F})$ and fix the mode of convergence that we use on this space.

Definition 3.1.12. A sequence $(\mathbb{P}_n)_{n\in\mathbb{N}}$ in $\mathcal{M}(\mathcal{D})$ converges to $\mathbb{P}$ in the sense of finite dimensional distributions if and only if $\mathbb{E}_{\mathbb{P}_n}[Z] \to \mathbb{E}_{\mathbb{P}}[Z]$ for every random variable Z of the form

$$Z = f(X_{t_1}, \dots, X_{t_m}),$$

for $t_1, \dots, t_m \in \mathbb{R}^+$ and continuous bounded $f\colon \mathbb{R}^m \to \mathbb{R}$.

As usual we denote by $\mathbb{R}^S$, where $S \subseteq \mathbb{R}^+$, the set of real valued functions on S. We equip $\mathbb{R}^S$ with the topology of pointwise convergence and denote by $\mathcal{F}^S$ it Borel σ-algebra. The coordinate process on $\mathbb{R}^S$ is denoted by X^S_t with natural filtration $(\mathcal{F}^S_t)_{t\geq 0}$, $\mathcal{F}^S_t = \sigma(X^S_s; s \in S, s \leq t)$.

We further equip the space of probability measures $\mathcal{M}\left((\mathbb{R}^S, F^S)\right)$ with the weak topology induced by the mappings $\mathbb{P} \to \mathbb{E}_{\mathbb{P}}[f]$ for all bounded continuous real functions f on $\mathbb{R}^S$. And for any $\mathbb{P}$ on $\mathcal{D}[0,\infty)$ we denote by $\mathbb{P}^S$ the probability measure on $(\mathbb{R}^S, \mathcal{F}^S)$ which is given through the restriction of the law of the coordinate process (X_t) under $\mathbb{P}$ to $t \in S$.

In other words, a sequence $(\mathbb{P}_n)$ of probability measures converges to $\mathbb{P}$ in the sense of f.d.d.s if and only if for all finite sets $S \in \mathbb{R}^+$ the sequence $\mathbb{P}_n^S$ converges weakly to $\mathbb{P}^S$.
On CP we use the topology of pointwise convergence, i.e. $C_n \to C$ if and only if $C_n(t,x) \to C(t,x)$ for all $(t,x) \in \mathbb{R}^+ \times \mathbb{R}$.

The first main result of this section reads as follows.

Theorem 45. *(Lowther 2008, cf. [Low08b, Theorem 1.3].) For any $C \in CP$ there exists a unique measure $\mathbb{P}$ on $(\mathcal{D}[0,\infty), \mathcal{F})$ under which X is an ACD martingale and $C(t,x) = \mathbb{E}[(X_t - x)^+]$.*

Proof. (Very rough sketch.) The existence part contains two steps. The idea and first part is to consider limits of processes that match the marginals at finite sets of time. Existence of the limit is shown via a tightness argument. The second part is to show that there in fact exist such sequences of processes, i.e. sequences of ACD martingale measures which match the given family of marginals at finite sets of time.

The fundamental Lemma 3.1 in [Low08b] shows, for $C \in CP$ and a sequence $(\mathbb{P}_n)_{n\in\mathbb{N}}$ of martingale measures on $(\mathcal{D}[0,\infty), \mathcal{F})$ such that $\mathbb{E}_{\mathbb{P}_n}[(X_t - x)^+] \to C(t,x)$, that there exists a subsequence $\mathbb{P}_{n_k}$ and a martingale measure $\mathbb{P}$ on $(\mathcal{D}[0,\infty), \mathcal{F})$ such that $\mathbb{P}_{n_k} \to \mathbb{P}$ in the sense of finite dimensional distributions. Furthermore the coordinate process X is a martingale under the limit measure $\mathbb{P}$, continuous in probability and satisfies $\mathbb{E}_{\mathbb{P}}[(X_t - x)^+] = C(t,x)$. [Low07, Corollary 1.3] then yields that X in fact is an ACD martingale.

The idea of the proof of the lemma is to consider a countable dense subset $S \subset \mathbb{R}^+$. This implies that $\mathbb{R}^S$ is polish. Furthermore, for a sequence $(\mathbb{P}_n)$ of martingale measures s.t. $\mathbb{E}_{\mathbb{P}_n}[(X_t - x)^+] \to C(t,x)$ the sequence $(\mathbb{P}_n^S)$ is tight and (at least a subsequence) converges weakly to a probability measure $\mathbb{Q}$ on $(\mathbb{R}^S, \mathcal{F}^S)$. Then one shows that (X_t^S) is a $\mathbb{Q}$-martingale to extend X_t^S to all $t \in \mathbb{R}^+$. The proof that (X_t^S) is continuous in probability is based on the fact that for a martingale the left and right limits X_{t-}^S and X_{t+}^S exist almost sure. Since X^S is right continuous in probability it has a cadlag version which implies that there exists a measure $\mathbb{P}$ on $\mathcal{D}[0,\infty)$ s.t. $\mathbb{P}^S = \mathbb{Q}$. X however is a martingale and continuous in probability under $\mathbb{P}$, so we can take limits of $t \in S$ to show that $\mathbb{E}_{\mathbb{P}}[(X_t - x)^+] = C(t,x)$. The convergence in the sense of f.d.d.s is shown via a standard argument.

What remains is to find an appropriate sequence $(\mathbb{P}_n)$ which matches the marginals at finitely many times. To this end one first constructs via Skorohod embeddding an ACD martingale X which matches the family of marginals at two times, $t_0 < t_1 \in \mathbb{R}^+$. I.e., $\mathbb{E}[(X_t - x)^+] = C(t,x)$ holds for $C \in CP$ and $t = t_0$ and $t = t_1$ (cf. Section 2.3 and [Low08a], p. 11 for details). Then this is extended to finitely many times, i.e., one constructs an ACD martingale measure $\mathbb{P}$ on $(\mathcal{D}[0,\infty), \mathcal{F})$ such that $\mathbb{E}_{\mathbb{P}}[(X_t - x)^+] = C(t,x)$ for all $t \in S$, $S \subset \mathbb{R}^+$ and $|S| < \infty$. So we can define an approximating sequence $(\mathbb{P}_n)$ s.t.

$$\mathbb{E}_{\mathbb{P}_n}[(X_{k/n} - x)^+] = C(k/n, x), \text{ for } k = 0, 1, \dots, n.$$

This implies that

$$\mathbb{E}_{\mathbb{P}_n}[(X_t - x)^+] \to C(t,x),$$

and the existence of the ACD martingale measure follows from the above cited fundamental lemma.

Uniqueness in the sense of f.d.d.s follows from the generalized backward equation developed in the previous section, which allows one to pass from equality in the one dimensional marginals to equality in all finite dimensional distributions. □

Given $C \in CP$ we will denote by $\mathbb{P}_C$ the unique ACD martingale measure that matches the family of marginals given via C. The map from the set of marginals to the set of martingale measures, as noted, is continuous.

Theorem 46. *(Lowther 2008, cf. [Low08b, Theorem 1.4].) Let $\mathbb{P}_C$ be the unique ACD martingale measure associated to $C \in CP$. Then the map*

$$CP \to \mathcal{M}(\mathcal{D}), \quad C \mapsto \mathbb{P}_C,$$

is continuous under pointwise convergence on CP and convergence in the sense of finite dimensional distributions on $\mathcal{M}(\mathcal{D})$. I.e, given any sequence $C_n \in CP$ converging pointwise to $C \in CP$ then $\mathbb{E}_{\mathbb{P}_{C_n}}[Z] \to \mathbb{E}_{\mathbb{P}_C}[Z]$ for every Z of the form (3.1.12).

Proof. Cf. [Low08b, Section 4] □

Remark 3.1.13. In particular, if we are given a (weak or strong) solution of a stochastic differential equation of the form

$$dX_t = \sigma(t, X_t)\, dB_t, \quad X_0 = x,$$

where σ is continuous bounded and measurable, then the solution is a continuous strong Markov martingale. So, in the case of a real valued diffusion we obtain from Theorem 44 that if two solutions are equal in the 1-d marginals they are already versions of each other. I.e., we now see why the approach of Dupire (cf. the introduction) was bound to succeed.

3.2 Itō-processes

Up to now we were concerned just with fitting martingales to marginals resp. with mimicking a certain martingale (Brownian motion) via other processes having the same marginals.

In this section we enlarge the scope of processes to be mimicked to (continuous) Itō-processes and discuss a result of Gyöngy. As mentioned, [Gyö86, Theorem 4.6.] roughly says that any time-inhomogeneous Itō-process can be mimicked by a Markov process which is obtained as the solution of a specific simpler SDE. The section is organized as follows: first we state the theorem, repeat Gyöngy's motivating derivation of a measure valued evolution equation, then sketch in some detail Gyöngy's proof and finally propose a somewhat simpler proof of a special case.

3.2.1 The Gyöngy setting

Theorem 47. *(Gyöngy 1986, cf. [Gyö86, Theorem 4.6].) Let*

$$d\xi(t, \omega) = \delta(t, \omega)\, dB_t + \beta(t, \omega)\, dt \tag{3.2.1}$$

be a multidimensional Itō-process, where δ and β are bounded, progressively measurable processes, s.t. $\delta\delta^T$ is uniformly positive definite. I.e., there is some $\lambda > 0$, *s.t.*

$$\sum_{i,j}^{n} \delta\delta^T(t,\omega)_{ij} x_i x_j \geq \lambda |x|^2 \quad \forall (t,\omega) \in [0,\infty) \times \Omega$$

and for every $x \in \mathbb{R}^n$*. Then there exists an SDE*

$$dX_t = a^{\frac{1}{2}}(t, X_t)\, dB_t + b(t, X_t)\, dt, \quad X_0 = 0, \tag{3.2.2}$$

which admits a weak solution having the same one-dimensional marginals as $\xi(t,\omega)$*. The coefficients are given by*

$$a(t,x) = \mathbb{E}[\delta\delta^T(t,\omega) |\, \xi(t,\omega) = x], \tag{3.2.3}$$

$$b(t,x) = \mathbb{E}[\beta(t,\omega) |\, \xi(t,\omega) = x]. \tag{3.2.4}$$

The proof of this theorem is rather involved, so we first repeat Gyöngy's motivation and show at which point one could take another way to prove a special case of the above result.

First, consider the process without drift

$$d\xi(t,\omega) = \delta(t,\omega)\, dB_t, \quad \xi(0) = 0. \tag{3.2.5}$$

Let $u(t,x)$ be a real valued smooth function that vanishes at infinity and apply Itō's formula. (Setting the drift term $\beta \equiv 0$ is no loss of generality. If the diffusion-term is uniformly elliptic, the drift term has no influence on existence or uniqueness of a solution of an SDE. Cf. [Bas98, Theorem 3.1.], p.130.)

$$\begin{aligned} du(t,\xi(t)) = \frac{\partial}{\partial t} u(t,\xi(t))\, dt &+ \sum_{i=1}^{n} \frac{\partial}{\partial x_i} u(t,\xi(t))\, d\xi_i(t) \\ &+ \frac{1}{2} \sum_{i,j=1}^{n} \delta\delta^T_{ij}(t,\omega) \frac{\partial^2}{\partial x_i \partial x_j} u(t, (\xi(t))\, d\langle \xi_i(t), \xi_j(t) \rangle. \end{aligned}$$

Using Fubini and the known orthogonality relations of multidimensional Brownian

Motion we get, as in Proposition 1.7.3,

$$\begin{aligned}
&\mathbb{E}\left[u(t,\xi(t)) - u(0,0)\right] = \\
&= \mathbb{E}\left[\int_0^t \left(\frac{\partial}{\partial s}u(s,\xi(s)) + \frac{1}{2}\sum_{i,j=1}^n \delta\delta^T_{ij}(s,\omega)\frac{\partial^2}{\partial x_i \partial x_j}u(s,(\xi(s))\right) ds\right] = \\
&= \int_0^t \mathbb{E}\left[\left(\frac{\partial}{\partial s} + \frac{1}{2}\sum_{i,j=1}^n \mathbb{E}[\delta\delta^T(s,\omega)|\xi(s)]_{ij}\frac{\partial^2}{\partial x_i \partial x_j}\right) u(s,\xi(s))\right] ds.
\end{aligned}$$

Rewriting the expectation in terms of the distribution of $\xi(t)$, $\mu_t(dx) = \mathbb{P}(\xi(t) \in dx)$, and using that the processes' starting at 0 is equivalent to an initial distribution $\mu_0(F) = \delta_0(F)$ for all F in $\mathcal{B}(\mathbb{R}^n)$ i.e. the Dirac measure concentrated at zero, the equation reads

$$\begin{aligned}
&\int\limits_{\mathbb{R}^n} u(t,x)\,d\mu_t(x) - \int\limits_{\mathbb{R}^n} u(0,x)\,d\delta_0(x) = \\
&= \int_0^t \left(\int_{\mathbb{R}^n}\left(\frac{\partial}{\partial s} + \frac{1}{2}\sum_{i,j=1}^n a_{ij}(s,x)\frac{\partial^2}{\partial x_i \partial x_j}\right) u(s,x)\,d\mu_s(x)\right) ds. \qquad (3.2.6)
\end{aligned}$$

Now consider $u(t,\widehat{X_t})$, where $(\widehat{X_t})$ is a weak solution of

$$dX_t = a^{\frac{1}{2}}(t,X_t)dB_t, \quad X(0) = 0, \qquad (3.2.7)$$

and apply the Itō-formula. It is clear that the distributions $\mathbb{P}(\widehat{X_t} \in dx)$ also satisfy equation (3.2.6) for every u in $\mathcal{C}_0^\infty([0,\infty)\times\mathbb{R}^n)$. According to Gyöngy, $a(t,x)$ can be chosen uniformly positive definite and bounded, hence one knows e.g. from [SV79] or [Bas98, Thm. VI.1.3], p.133, that the above SDE has a weak solution, which means that, if equation (3.2.6) would suffice to identify the distributions, we would be done. And indeed we will argue along these lines in proving the above mentioned special case of Theorem 47, since equation (3.2.6) can be viewed as a measure valued evolution equation of the form

$$\int_E f d\nu_t = \int_E f d\nu_0 + \int_0^t \left(\int_E \mathcal{L} f d\nu_s\right) ds,\ f \in D(\mathcal{L}), \qquad (3.2.8)$$

where E denotes the (polish) state space and $D(\mathcal{L})$ the domain of $\mathcal{L}$. In our case

$$\begin{aligned}\mathcal{L} &= \frac{\partial}{\partial t} + \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(s,x)\frac{\partial^2}{\partial x_i \partial x_j} = \\ &= \frac{\partial}{\partial t} + \mathcal{L}_t\end{aligned}$$

is the generator of the (mimicking) diffusion (3.2.7). A Theorem of Bhatt and Karandikar states that, if the martingale problem for an operator $\mathcal{L}$ is well posed and if two families of probability measures (μ_t) and (ν_t) satisfy the evolution equation (3.2.8), then they are already the same. In our setting this means: if the martingale problem for the operator $\frac{\partial}{\partial t} + \mathcal{L}_t$ associated to the SDE (3.2.7) is well posed, then equation (3.2.6) indeed identifies the one-dimensional marginals of (ξ_t) and $(\widehat{X}_t)$.

However, within the broad assumptions of Theorem 47 this does not work, since in general the martingale problem is not well posed for bounded and measurable coefficients, except in dimension one. Moreover, Gyöngy asserts that uniqueness results via such measure valued equations have only been proved for smooth coefficients $a(t,x)$. But, since uniqueness of the measure valued equation is directly related to well posedness of the martingale problem for the operator $\mathcal{L}$, it is clear from Theorem 22 that uniqueness holds under much broader assumptions then smoothness.

We now sketch the construction of Gyöngy and note that he does not consider the associated operator, but deals with the stochastic differential equation

$$dX_t = a^{\frac{1}{2}}(t, X_t)\, dB_t + b(t, X_t)\, dt$$

itself, where $a(t,x) = \mathbb{E}[\delta\delta^T(t,\omega)|\xi(t,\omega) = x]$ and b accordingly.

Remark 3.2.1. If we want to use Itō's classical uniqueness and existence theorem (cf. Theorem 18 above), the diffusion matrix $a(t,x)$ has to admit a Lipschitz continuous square root $a^{\frac{1}{2}}$, i.e. a matrix $(\sigma_{ij}(t,x))$ which is Lipschitz in the sense of Theorem 18 and such that $\sigma\sigma^T = a$. A theorem by Freidlin [Fre68] ensures that such a factorization exists for symmetric nonnegative definite matrices (c_{ij}), the entries $c_{ij}(x)$ of which are in $C^2(\mathbb{R}^n)$.

However, since $\delta(t,\omega)$ in Theorem 47 is assumed just to be bounded and measurable, $a(t,x) = \mathbb{E}[\delta\delta^T(t,\omega)|\xi(t,\omega) = x]$ is not differentiable in general and more efforts are necessary. There are two key elements in the construction of Gyöngy.

(i) An approximation procedure using smoothed coefficients a_{ij}^{ε}, originally due to N.V. Krylov, in order to get strong solutions X^{ε} for the smoothed equation.
(ii) The Green measure χ of a process $(X_t)_{t\geq 0}$, which measures the time the process spends in a set $\Gamma \in \mathcal{B}(\mathbb{R}^n)$ (before being killed with rate γ, where $\gamma(t,\omega)$ is nonnegative and adapted).

$$\chi(\Gamma) := \mathbb{E}\left[\int_0^{\infty} \mathbb{1}_{\Gamma}(X_t) \exp\left(-\int_0^t \gamma(s,\omega)\, ds\right) dt\right]$$

The idea of Gyöngy's proof consists roughly in showing, that the Green measures of the time-space solutions $(u^{\varepsilon}(t), X_t^{\varepsilon})$ of the smoothed equation converge weakly to the Green measure of (t, ξ_t), i.e. to $\mathbb{P}(\xi_t \in dx)dt$. Furthermore, the distribution of $(u^{\varepsilon}, X^{\varepsilon})$ on the path space $\mathcal{C}[0,T]$ converges weakly to the distribution of (t, X_t), where X_t is a weak solution of (3.2.2). Finally one notes that the coincidence of the Green measures of the time-space process (t, X_t) and (t, ξ_t) implies the coincidence of the one-dimensional distributions of the processes X and ξ.

In the following we won't repeat Gyöngy's proof in detail, but sketch the (rather technical) construction as comprehensibly as possible. First, we simplify equation (3.2.6) by letting $t \to \infty$ which cancels the first term since $u(t,x)$ vanishes at infinity. We obtain

$$\int_0^{\infty}\int_{\mathbb{R}^n} u(s,x)\, d\delta_{(0,0)}(ds,dx) = -\int_0^{\infty}\int_{\mathbb{R}^n} \left(\frac{\partial}{\partial s} + \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(s,x)\frac{\partial^2}{\partial x_i \partial x_j}\right) u(s,x)\, d\mu_s(x)\, ds, \tag{3.2.9}$$

where $\delta_{(0,0)}$ is the Dirac measure at $(0,0) \in \mathbb{R}^+ \times \mathbb{R}^n$.

Starting from this we consider an inequality satisfied by arbitrary measures ν and μ on $[0,\infty) \times \mathbb{R}^n$ (instead of $d\delta_{(0,0)}$ and $d\mu_s(x)ds$),

$$\int_0^{\infty}\int_{\mathbb{R}^n} u(t,x)\, \nu(dt,dx) \geq -\int_0^{\infty}\int_{\mathbb{R}^n} \mathcal{L}u(t,x)\, \mu(dt,dx),$$

which is valid for every nonnegative function $u \in \mathcal{C}_0^{\infty}([0,\infty) \times \mathbb{R}^n)$ and clearly covers the case of equation (3.2.6). Now we smooth the coefficients of the operator $\mathcal{L}$ via Radon-Nikodym derivatives.

The smoothing of bounded measurable functions g proceeds in two steps. First, for a measure μ one defines the measure $g\mu$ being the measure with density g w.r.t. μ. Then, as smoothing device, one fixes a nonnegative function $k \in \mathcal{C}^{\infty}(\mathbb{R}^n)$ such that

$\int k(x)\,dx = 1$ and $k(-x) = k(x)$ for all x and defines

$$k^{\varepsilon}(x) = \frac{1}{\varepsilon^n} k\left(\frac{x}{\varepsilon}\right) \qquad \text{for every } \varepsilon > 0.$$

This induces a measure $k^{(\varepsilon)}(x)\,dx$, which we convolute with the measure $g\mu$ in order to define a measure $(g\mu)^{(\varepsilon)}(x)\,dx$ with density function

$$(g\mu)^{(\varepsilon)}(x) = \int k^{\varepsilon}(x-y)\,g(y)\,\mu(dy)$$

w.r.t. the Lebesgue measure. If $\mu^{(\varepsilon)}(x)$ is positive on $\mathbb{R}^n$ for all $\varepsilon > 0$, then the measure $(g\mu)^{(\varepsilon)}(x)\,dx$ has a smooth Radon-Nikodym derivative w.r.t. to $\mu^{(\varepsilon)}(x)\,dx$ for every bounded measurable g, which is precisely the desired smoothed function

$$g_{(\varepsilon)} = \frac{(g\mu)^{(\varepsilon)}}{\mu^{(\varepsilon)}}. \tag{3.2.10}$$

In the case of time-homogeneous processes now we are able to smooth the bounded measurable coefficients a_{ij}, b_i getting $a^{ij}_{(\varepsilon)}$, $b^i_{(\varepsilon)}$ accordingly. In the case of time dependent coefficients, instead of k one considers smoothing kernels $h(t,x) = \psi(t)k(x)$, where $\psi(t) \in \mathcal{C}_0^\infty(\mathbb{R})$ is such that $\int \psi(t)\,dt = 1$, $\psi(t) > 0$ on $(-1,0)$ and zero elsewhere. The smoothing is defined analogously. Finally we consider the operator associated to (3.2.2) (with an additional killing term c),

$$\mathcal{L} = \frac{\partial}{\partial t} + \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(t,x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(t,x)\frac{\partial}{\partial x_i} - c(t,x), \tag{3.2.11}$$

and the family of smoothed operators

$$\mathcal{L}_\varepsilon = \frac{\partial}{\partial t} + \frac{1}{2}\sum_{i,j=1}^{n} a^{ij}_{(\varepsilon)}(t,x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b^i_{(\varepsilon)}(t,x)\frac{\partial}{\partial x_i} - c_{(\varepsilon)}(t,x). \tag{3.2.12}$$

Accordingly the stochastic differential equations under consideration read

$$dX_t = (a_{(\varepsilon)})^{\frac{1}{2}}(t+s, X_t)\,dB_t + b_{(\varepsilon)}(t+s, X_t)\,dt \tag{3.2.13}$$

for every $s \geq 0$.

We now state (without proof) the Fundamental Lemma, which gives an estimate for the Green measures of solutions to (3.2.13) via some sort of resolvent operator. Later on this estimate is used to prove weak convergence of the Green measures of (strong) solutions of (3.2.13).

Lemma 3.2.2. *(Cf. [Gyö86, Lemma 2.2.], p.506) Assume that* $0 < \mu^{(\varepsilon)}(t,x) < \infty$ *on* $(0,\infty) \times \mathbb{R}^n$ *for every* $\varepsilon > 0$*. Suppose that*

$$\int\limits_{[0,\infty)\times\mathbb{R}^n} u(t,x)\,\nu(dt,dx) \geq -\int\limits_{[0,\infty)\times\mathbb{R}^n} \mathcal{L}u(t,x)\,\mu(dt,dx) \tag{3.2.14}$$

for every $u \in \mathcal{C}_0^\infty([0,\infty)\times\mathbb{R}^n)$*.*
Let $X_\varepsilon^{s,x}$ *be a solution of equation* (3.2.13) *with initial condition* $X_\varepsilon^{s,x}(0) = x \in \mathbb{R}^n$*, and define*

$$\mathfrak{R}^\varepsilon f(s,x) = \mathbb{E}\left[\int_0^\infty f(s+t, X_\varepsilon^{s,x}(t)) \exp\left(-\int_0^t c_{(\varepsilon)}(s+r, X_\varepsilon^{s,x}(r)\,dr\right) dt\right].$$

Then for every nonnegative $f \in \mathcal{C}_0^\infty([0,\infty)\times\mathbb{R}^n)$

$$\int\limits_{[0,\infty)\times\mathbb{R}^n} (\mathfrak{R}^\varepsilon f)^{(\varepsilon)}\,\nu(dt,dx) \geq \int\limits_{[0,\infty)\times\mathbb{R}^n} f^{(\varepsilon)}\,\mu(dt,dx) \tag{3.2.15}$$

for every $\varepsilon > 0$*, and thus*

$$\liminf_{\varepsilon\searrow 0} \int (\mathfrak{R}^\varepsilon f)^{(\varepsilon)}\,\nu(dt,dx) \geq \int f\,\mu(dt,dx). \tag{3.2.16}$$

If the inequality (3.2.14) *is reversed, then it is also reversed in* (3.2.15) *and* (3.2.16) *with* lim sup *instead of* lim inf*.*

Proof. Cf. [Gyö86], p.509ff. □

Before we state and prove the convergence of the Green measures, we sketch the construction of the coefficients a_{ij} and b_i which again are defined as Radon-Nikodym derivatives. Precisely this construction ensures the applicability of the Fundamental Lemma 3.2.2 and also the convergence

$$a_{(\varepsilon)}(t,x) \to a(t,x) = \mathbb{E}[\delta\delta^T(t,\omega)|\xi(t,\omega) = x].$$

Recall our given process

$$\xi(t,\omega) = \delta(t,\omega)\,dB_t + \beta(t,\omega)\,dt.$$

Set $\alpha = \delta\delta^T$, $\eta(t) = (t, \xi(t))$ and let γ denote the killing rate of the Green measure μ of the time-space process η.

$$\mu(\Gamma) = \mathbb{E}\left[\int_0^\infty \mathbb{1}_\Gamma(\eta(t)) \exp(-\varphi(t))\, dt\right],$$

for every $\Gamma \in \mathcal{B}([0,\infty) \times \mathbb{R}^n)$, where

$$\varphi(t) = \int_0^t \gamma(s)\, ds.$$

Now we observe that for bounded measurable nonnegative processes $\phi(t,\omega)$ the measure

$$\mu_\phi(\Gamma) = \mathbb{E}\left[\int_0^\infty \mathbb{1}_\Gamma(\eta(t))\phi(t,\omega) \exp(-\varphi(t))\, dt\right],$$

is absolutely continuous w.r.t. μ. I.e., we are able to define the coefficient functions $a = (a^{ij}(t,x))$, $b = (b^i(t,x))$ and $c = c(t,x)$ as Radon-Nikodym derivatives w.r.t. the Green measure μ of $\eta(t) = (t, \xi(t))$:

$$a^{ij} := \frac{d\mu_\alpha^{ij}}{d\mu}, \qquad b^i := \frac{d\mu_\beta^i}{d\mu} \quad \text{and } c := \frac{d\mu_\gamma}{d\mu}.$$

The so defined Borel measurable coefficient functions still have to be smoothed according to equation (3.2.10), using the kernel $h(t,x)$ and supposing that $0 < \mu^{(\varepsilon)} < \infty$ on $[0,\infty) \times \mathbb{R}^n$ for every $\varepsilon > 0$. Hence we get the smoothed coefficient functions

$$a_{(\varepsilon)} = (a^{ij}_{(\varepsilon)}), \qquad b_{(\varepsilon)} = (b^i_{(\varepsilon)}), \qquad c_{(\varepsilon)},$$

and are able to state and prove the convergence result.

Lemma 3.2.3. *(Cf. [Gyö86, Lemma 4.1.]) Let $z_\varepsilon(t) = (u_\varepsilon(t), X_\varepsilon(t))$ be the solution of the equation*

$$\begin{aligned} du_\varepsilon(t) &= dt \\ dX_\varepsilon(t) &= (a_{(\varepsilon)})^{\frac{1}{2}}(u_\varepsilon(t), X_\varepsilon(t))\, dB_t + b_{(\varepsilon)}(u_\varepsilon(t), X_\varepsilon(t))\, dt \\ z_\varepsilon(0) &= z_0^\varepsilon, \end{aligned} \tag{3.2.17}$$

where z_0^ε is a r.v. in $\mathbb{R}^{n+1}$ independent of the Brownian motion B and having density $h^{(\varepsilon)}$. Then for $\varepsilon \searrow 0$ the Green measure of the process z_ε, with killing rate $c_{(\varepsilon)}(z_\varepsilon(t))$, converges weakly to the Green measure μ.

Proof. Let ν be the Dirac measure concentrated at $0 \in \mathbb{R}^{n+1}$. Applying the Itō-formula to $u(t,\xi(t))$ we get, for every $u \in \mathcal{C}_0^\infty([0,\infty)\times\mathbb{R}^n)$,

$$\begin{aligned}
\int u(t,x)\,\nu(dt,dx) = u(0) = \\
&= -\mathbb{E}\left[\int_0^\infty \left(\frac{\partial}{\partial t}u(t,\xi_t) + \frac{1}{2}\sum_{i,j}\alpha^{ij}(t)\frac{\partial^2}{\partial x_i \partial x_j}u(t,\xi_t) + \right.\right.\\
&\quad \left.\left. + \sum_i \beta^i(t)\frac{\partial}{\partial x_i}u(t,\xi_t) - \gamma(t)u(t,\xi(t))\right)\exp(-\varphi(t))\,dt\right] = \\
&= -\int_{[0,\infty)\times\mathbb{R}^n} \frac{\partial}{\partial t}u(t,x)\,\mu(dt,dx) + \frac{1}{2}\sum_{i,j}\frac{\partial^2}{\partial x_i\partial x_j}u(t,x)\,\mu_\alpha^{ij}(dt,dx) + \\
&\quad + \sum_i \frac{\partial}{\partial x_i}u(t,x)\,\mu_\beta^i(dt,dx) - u(t,x)\,\mu_\gamma(dt,dx) = \\
&= -\int_{[0,\infty)\times\mathbb{R}^n} \mathcal{L}u(t,x)\,\mu(dt,dx),
\end{aligned}$$

where the last step is due to the definition of the coefficients of $\mathcal{L}$ as Radon-Nikodym derivatives w.r.t. to μ. From the Fundamental Lemma 3.2.2 we get for every $f \in \mathcal{C}_0^\infty([0,\infty)\times\mathbb{R}^n)$

$$\int_{[0,\infty)\times\mathbb{R}^n} f(t,x)\,\mu(dt,dx) = \lim_{\varepsilon\searrow 0}(\mathfrak{R}^\varepsilon f)^{(\varepsilon)}(0).$$

And we see that weak convergence of the Green measures holds, since $(\mathfrak{R}^\varepsilon f)^{(\varepsilon)}$ is the integral of f on $[0,\infty)\times\mathbb{R}^n$ w.r.t. the Green measure of $z_\varepsilon(t)$. □

Remark 3.2.4. In order to prove Theorem 47 yet some intermediary results are missing. However, all the ancillary propositions in [Gyö86] are more or less clear from the construction. We just list the necessary results.

(i) The Green measure μ is equivalent to the Lebesgue measure on $[0,\infty)\times\mathbb{R}^n$.

(ii) For $\varepsilon \to 0$:
   a) $a_{(\varepsilon)}(t,x) \to a(t,x)$
   b) $b_{(\varepsilon)}(t,x) \to b(t,x)$
   c) $c_{(\varepsilon)}(t,x) \to c(t,x)$

   for Lebesgue almost all $(t,x) \in [0,\infty)\times\mathbb{R}^n$.

(iii) For $\gamma(t)$ deterministic, we get
   a) $a^{ij}(t,x) = \mathbb{E}[\alpha^{ij}(t,\omega)|\xi(t,\omega) = x]$
   b) $b^i(t,x) = \mathbb{E}[\beta^i(t,\omega)|\xi(t,\omega) = x]$
   c) $c(t,x) = \mathbb{E}[\gamma(t)|\xi(t) = x]$

for Lebesgue almost all $(t,x) \in [0,\infty) \times \mathbb{R}^n$.

(iv) For a sequence (ε_n), $\varepsilon_n \to 0$, the distribution of the process z_ε on $C[0,T]$ converges weakly to the distribution of the process $z(t) = (t, X_t)$, where (X_t) is a weak solution of (3.2.2) with initial condition $X_0 = 0 \in \mathbb{R}^n$.

Proof of Theorem 47. (Cf. [Gyö86, Theorem 4.6.], p.516)
From Lemma 3.2.3 and Remark 3.2.4 we get that the Green measure of (t,ξ_t) (with killing rate γ is identical to the Green measure of (t, X_t) (with killing rate $c(t, X_t)$). Take $\gamma(t) \equiv 1$. Then clearly $c(t,x) = \gamma = 1$ and

$$\mathbb{E}\left[\int_0^\infty e^{-t} f(t,\xi_t)\,dt\right] = \mathbb{E}\left[e^{-t} f(t,X_t)\,dt\right]$$

for every bounded nonnegative Borel measurable function f. Taking $f(t,x) = e^{-\lambda t} g(x)$ with arbitraty $\lambda \geq 0$ and $g \in C_0(\mathbb{R}^n)$, we get

$$\int_0^\infty e^{-\lambda t} e^{-t} \mathbb{E}[g(\xi_t)]\,dt = \int_0^\infty e^{-\lambda t} e^{-t} \mathbb{E}[g(X_t)]\,dt$$

for every $\lambda \geq 0$ and $g \in C_0(\mathbb{R}^n)$.
Since $\mathbb{E}[g(\xi_t)]$ and $\mathbb{E}[g(X_t)]$ are continuous in t we obtain from the above equation that for every $t \geq 0$ and every $g \in C_0(\mathbb{R}^n)$

$$\mathbb{E}[g(\xi_t)] = \mathbb{E}[g(X_t)].$$

Hence the distributions of ξ_t and X_t are the same for all $t \geq 0$. □

3.2.2 Proof via evolution equations

As we have seen, the construction of Gyöngy (even if summarized) is rather demanding. In the following we sketch another approach to the problem which yields a proof of Theorem 47 if we additionally require the diffusion matrix $a(t,x) = \mathbb{E}[\delta\delta^T(t,\omega)|\xi(t,\omega) = x]$ to be continuous.

The proposed proof, as mentioned, is not based on the Itō-theory of SDEs, but on the martingale problem approach by Stroock and Varadhan. I.e. we are not concerned with equation (3.2.2), but with the associated operator. The advantage of this approach lies in the fact that one can use the diffusion coefficient a of the operator $\mathcal{L}$ directly, i.e., we do not have to factorize the positive definite matrix $a(t,x)$ into

$a^{\frac{1}{2}}(t,x)$. In particular, the hardest part of the Gyöngy-proof, the smoothing of the coefficients a and b is no longer required.

Put in a nutshell, the proposed proof is based on two results.

(i) Theorem 22, which gives sufficient conditions for existence and uniqueness of a solution to the martingale problem for an operator $\mathcal{L}$.

(ii) [BK93, Theorem 5.2.], which ensures that the solution $\{\nu_t\}_{t\geq 0}$, $\nu_t \in \mathcal{M}(\mathbb{R}^n)$, of a measure valued evolution equation of the form

$$\int_{\mathbb{R}^n} f d\nu_t = \int_{\mathbb{R}^n} f d\nu_0 + \int_0^t \left(\int_{\mathbb{R}^n} (\frac{\partial}{\partial s} + \mathcal{L}_s) f d\nu_s \right) ds, \quad f \in D(\mathcal{L}), \qquad (3.2.18)$$

is unique if the martingale problem for $\mathcal{L} = \frac{\partial}{\partial t} + \mathcal{L}_t$ is well posed and if $t \mapsto \nu_t(U)$ is measurable for every Borel set U in $\mathbb{R}^n$.

Note that, if X is a solution of (3.2.2) or a solution of the martingale problem for the associated operator $\mathcal{L}$ and initial distribution ν_0, then the family of one dimensional marginals $\mu_t(dx) = \mathbb{P}(X_t \in dx)$ is a solution to the above equation (cf. equation (3.2.6)).

Proof of Theorem 47, special case. Assume in addition to the assumptions of Theorem 47 that the diffusion matrix $a(t,x) = \mathbb{E}[\delta\delta^T(t,\omega)|\xi(t,\omega) = x]$ is continuous. Then $a(t,x)$ and $b(t,x) = \mathbb{E}[\beta(t,\omega)|\xi(t,\omega) = x]$ satisfy the assumptions of Theorem 22, i.e. the martingale problem for

$$\mathcal{L} = \frac{\partial}{\partial t} + \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(t,x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(t,x) \frac{\partial}{\partial x_i}$$

is well posed for all $x \in \mathbb{R}^n$, hence for any initial distribution ν_0 on $\mathbb{R}^n$. Let $\widetilde{X}$ denote the unique solution w.r.t. the initial distribution δ_0, i.e. $\widetilde{X}_0 = 0$.

Furthermore, as is clear from the motivation at the beginning of the section (cf. (3.2.6)), both families of distributions $\{\mu_t(dx)\}_t = \{\mathbb{P}(\xi_t \in dx)\}_t$ and $\{\nu_t\}_t = \{\mathbb{P}(\widetilde{X}_t \in dx)\}_t$, $t \geq 0$, satisfy the measure valued equation (3.2.18) for $\mathcal{L}$ above. And, since we only consider processes with continuous paths, the maps $t \mapsto \xi_t(\omega)$ and $t \mapsto \widetilde{X}_t(\omega)$ are continuous, hence the maps $t \mapsto \mu_t(U)$ and $t \mapsto \nu_t(U)$ are measurable for all Borel sets U. In order to be able to apply [BK93, Theorem 5.2.] and to obtain uniqueness for the measure valued equation, the chosen domain $D(\mathcal{L})$ has to satisfy two separability properties. A measurablility property is required for the solutions of the martingale problem.

(i) $D(\mathcal{L})$ is an algebra that separates points and vanishes nowhere.
(ii) There exists a countable subset $\{g_k\} \subseteq D(\mathcal{L})$ such that

$$bp - closure\left(\{(g_k, \mathcal{L}g_k)\colon k \geq 1\}\right) \supseteq \{(g, \mathcal{L}g)\colon g \in D(\mathcal{L})\},$$

where the bp-closure is the closure under bounded-pointwise convergence.
(iii) Every progressively measurable solution to the martingale problem for $(\mathcal{L}, \nu_0)$ admits a càdlàg modification.

For compact metric state spaces E, assumption (iii) is always fulfilled. If E is locally compact and metric, (iii) holds if the operator $\mathcal{L}$ is conservative. I.e., if $\mathcal{L}$ is any second order differential operator on $\mathbb{R}^n$ with bounded coefficients, it is conservative, i.e. (iii) is satisfied in our case. Clearly the chosen domain $D(\mathcal{L}) = \mathcal{C}_0^\infty([0,\infty) \times \mathbb{R}^n)$ is an algebra that separates points and vanishes nowhere, hence property (i) follows. With regard to property (ii) one should think of the space of polynomials with rational coefficients to obtain the required countable dense subset.

Since both ξ and $\widetilde{X}$ start at zero, we get that the solution of the measure valued equation (3.2.18) is unique for $\mathcal{L}$ above and

$$\mu_t(dx) = \mathbb{P}(\xi_t \in dx) = \mathbb{P}(\widetilde{X}_t \in dx) = \nu_t(dx) \quad \text{for all } t \geq 0.$$

□

Remark 3.2.5. Gyöngy's result only yields the existence of a certain mimicking process. In general the solution of the SDE (3.2.2) is not unique. However, if the given Itō-process is real valued and has no drift then Theorem 44 yields that the solution of the constructed SDE (which clearly then is a continuous strong Markov martingale) is unique, i.e., the Markovian projection is well defined.

Remark 3.2.6. In the proof of the special case of Theorem 47 we solved the mimicking SDE with the help of the martingale problem. Moreover, the uniqueness of the solution of the martingale problem was crucial for the solution of the measure valued equation (3.2.6). So, by construction the mimicking process is unique and the Markovian projection is well defined. Note that in this case we obtain uniqueness even *with* drift. The projection is not restricted to martingales.

Remark 3.2.7. In dimension one and two the martingale problem is well posed if the coefficients of the operator are just bounded and measurable. So, our proof of the special case is valid without any additional assumption beyond the Gyöngy setting. In other words, in dimension one and two it is not a proof of a special case, but one of the general case. This means that for a real valued Itō-process the mimicking process is unique and the Markovian projection is well defined even with drift.

Bibliography

[Alb08] J. M. P. Albin. A continuous non-Brownian motion martingale with Brownian motion marginal distributions. *Statist. Probab. Lett.*, 78(6):682–686, 2008.

[Bas95] R. F. Bass. *Probabilistic techniques in analysis.* Probability and its Applications (New York). Springer-Verlag, New York, 1995.

[Bas98] R. F. Bass. *Diffusions and elliptic operators.* Probability and its Applications (New York). Springer-Verlag, New York, 1998.

[BK93] A. G. Bhatt and R. L. Karandikar. Invariant measures and evolution equations for Markov processes characterized via martingale problems. *Ann. Probab.*, 21(4):2246–2268, 1993.

[Cra36] H. Cramér. Über eine Eigenschaft der normalen Verteilungsfunktion. *Math. Z.*, 41(1):405–414, 1936.

[Dup97] B. Dupire. Pricing and hedging with smiles. In *Mathematics of derivative securities (Cambridge, 1995)*, volume 15 of *Publ. Newton Inst.*, pages 103–111. Cambridge Univ. Press, Cambridge, 1997.

[EK86] S. N. Ethier and T. G. Kurtz. *Markov processes.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. Characterization and convergence.

[Föl81] H. Föllmer. Calcul d'Itô sans probabilités. In *Seminar on Probability, XV (Univ. Strasbourg, Strasbourg, 1979/1980) (French)*, volume 850 of *Lecture Notes in Math.*, pages 143–150. Springer, Berlin, 1981.

[Fre68] M. I. Freĭdlin. The factorization of nonnegative definite matrices. *Teor. Verojatnost. i Primenen.*, 13:375–378, 1968.

[FWY99] H. Föllmer, C.-T. Wu, and M. Yor. Canonical decomposition of linear transformations of two independent Brownian motions motivated by models of insider trading. *Stochastic Process. Appl.*, 84(1):137–164, 1999.

[FWY00] H. Föllmer, C.-T. Wu, and M. Yor. On weak Brownian motions of arbitrary order. *Ann. Inst. H. Poincaré Probab. Statist.*, 36(4):447–487, 2000.

[Gyö86] I. Gyöngy. Mimicking the one-dimensional marginal distributions of processes having an Itô differential. *Probab. Theory Relat. Fields*, 71(4):501–516, 1986.

[HK06] K. Hamza and F. C. Klebaner. A Family of non-Gaussian Martingales with Gaussian Marginals. *ArXiv Mathematics e-prints*, April 2006.

[HK07] K. Hamza and F. C. Klebaner. A family of non-Gaussian martingales with Gaussian marginals. *J. Appl. Math. Stoch. Anal.*, pages Art. ID 92723, 19, 2007.

[Kel72] H. G. Kellerer. Markov-Komposition und eine Anwendung auf Martingale. *Math. Ann.*, 198:99–122, 1972.

[Kni81] F. B. Knight. *Essentials of Brownian motion and diffusion*, volume 18 of *Mathematical Surveys.* American Mathematical Society, Providence, R.I., 1981.

[KS88] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1988.

[Low07] G. Lowther. Limits Of One Dimensional Diffusions. *ArXiv e-prints*, December 2007.

[Low08a] G. Lowther. A Generalized Backward Equation For One Dimensional Processes. *ArXiv e-prints*, March 2008.

[Low08b] G. Lowther. Fitting Martingales To Given Marginals. *ArXiv e-prints*, August 2008.

[Low08c] G. Lowther. Nondifferentiable functions of one-dimensional semimartingales. *ArXiv e-prints*, February 2008.

[Low08d] G. Lowther. Properties of Expectations of Functions of Martingale Diffusions. *ArXiv e-prints*, January 2008.

[MY02] D. B. Madan and M. Yor. Making Markov martingales meet marginals: with explicit constructions. *Bernoulli*, 8(4):509–536, 2002.

[Øks98] B. Øksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, fifth edition, 1998. An introduction with applications.

[Ole08] K. Oleszkiewicz. On fake Brownian motions. *Statist. Probab. Lett.*, 78(11):1251–1254, 2008.

[Pro04] P. E. Protter. *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, second edition, 2004. Stochastic Modelling and Applied Probability.

[RY99] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[Sat99] Ken-iti Sato. *Lévy processes and infinitely divisible distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. Translated from the 1990 Japanese original, Revised by the author.

[Shr04] S. E. Shreve. *Stochastic calculus for finance. II*. Springer Finance. Springer-Verlag, New York, 2004. Continuous-time models.

[Str65] V. Strassen. The existence of probability measures with given marginals. *Ann. Math. Statist.*, 36:423–439, 1965.

[SV79] D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1979.

Curriculum Vitae

am 6. März 1979 in Linz als Friedrich Penkner geboren;
drei Geschwister (Stefan, Angela, Bernhard);
ein Vater (Friedrich Penkner), eine Mutter (Rosa-Maria Penkner);
aufgewachsen im Mühlviertel (Oberösterreich);
von 1982-1985 Kindergarten in Alberndorf;
von 1985-1989 Volksschule in Kaindorf;
von 1989-1993 Hauptschule in Altenberg bei Linz;
von 1993-1998 HTBLA II Linz, Paul-Hahnstraße;
Matura am 1.7.1998;
von 1998-1999 Zivildienst im Diakoniewerk Gallneukirchen;
von 1999-2004 Studium der Philosophie und Deutschen Philologie an der Universität Wien;
Abschluss mit Mag.phil. am 4.6.2004;
von 2004-2006 Doktoratsstudium der Philosophie;
Abschluss mit Dr.phil. am 10.1.2007;
von 2002-2010 Studium der Mathematik an der Universität Wien.